THREE ESSAYS IN AGRICULTURAL ECONOMICS: INTERNATIONAL TRADE, DEVELOPMENT AND COMMODITY PROMOTION

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

In the Department of Agricultural Economics

University of Saskatchewan

Saskatoon

By

RYAN TYLER CARDWELL

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Agricultural Economics

University of Saskatchewan

51 Campus Drive

Saskatoon, Saskatchewan

S7N 5A8

ABSTRACT

This thesis contains three essays on topics in agricultural economics. Essays one and two share a focus on international trade and economic development, and essays two and three apply dynamic tools to agricultural economic policy issues.

Essay one analyses trade-related implications of a developing country's decision to adopt genetically-modified crop technology. A fixed-proportions model is constructed that evaluates the welfare implications of a range of adoption policies and export market responses. The model in this essay illustrates the importance of the prospective adopter formulating a projection of probable export market effects *before* making an adoption decision and of the role that high transaction costs may play in a developing country's adoption decision. The model also considers the effects of a new policy tool; a check-off style levy on genetically-modified technology in place of a technology-use agreement. A levy could be useful tool in developing countries, which are characterised by high transaction costs.

Essay two models the effects of emergency food aid on a recipient country's agricultural industry. This essay formulates a definition of "needed" aid in the context of a food emergency and constructs an optimal control model that solves a path of aid shipments that best meets that need. The effects of a range of food aid paths on recipient-country agricultural production are illustrated through numerical simulations. There are two key results. First, a non-optimal amount of aid can hinder a recipient-

country's recovery from an exogenous food shock. Second, an exogenous shock can affect farmer revenue and therefore impact planting decisions. This effect must be considered in aid allocation policies.

Essay three uses time-series econometric techniques to develop a demand model that assesses the effectiveness of commodity advertising. This essay describes the importance of considering long-run and dynamic effects in demand systems, especially in the case of closely substitutable commodities. A demand system that tests for and accommodates dynamic and time-series properties is developed and applied to US meat data. The results of this model are compared to a traditional static demand system. The dynamic model produces econometrically and theoretically sound results and generates some more intuitively appealing estimates.

ACKNOWLEDGEMENTS

Thank you to my supervisor, William A. Kerr, for providing me every opportunity to become a successful researcher and for providing an example of how I hope to teach and engage students during my academic career.

I would also like to thank my advisory committee, Richard Gray, Ken Belcher, Peter Phillips and Kien Tran. The broad spectrum of talents of each of these advisors provided assistance with the very different components of this thesis. Thanks also to my external examiner, Richard Barichello, for helping to improve the quality of each of the three essays contained in this dissertation.

I am also appreciative of the faculty, staff and my fellow students in the Department of Agricultural Economics for making my experience at the University of Saskatchewan enjoyable.

I gratefully acknowledge the financial support of the Canadian Wheat Board, the College of Graduate Studies and Research and the trustees of the C. Paul W. and Marianne M. Ziehlke Bursary in Agriculture.

Finally, thanks to Karolina for making me smile and laugh through the last five years.

This thesis is dedicated to my parents, Sharon and Ted, and to my brothers, Stephen and Dean. This would not have been possible without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ESSAY 1: DEVELOPING COUNTRIES' DECISION TO ADOPT GENETICALLY-MODIFIED CROP TECHNOLOGY: FRAMING THE INTERNATIONAL TRADE POLICY IMPLICATIONS

1.1. Introduction

Developing countries may have the most to gain and the most to lose from adoption of genetically-modified crop technology. Genetically-modified crops that increase agricultural productivity have the potential to enhance developing countries' agricultural comparative advantages and reduce staple food prices. Genetically-modified technologies that contain enhanced nutritional characteristics have the potential to improve diets, and create crops that are hardier in the face of weather and pest shocks, thereby stabilizing food security. Some developing countries are quickly adopting genetically-modified crop technology and it appears as though consumers in developing countries are amenable to consuming foods that are produced from genetically-modified crops.[1]

[1] Some exceptions include Uganda, Zambia and Zimbabwe (Haggui).

There exist, however, several significant downside risks to the adoption of genetically-modified crop technology in developing countries.[2] These risks involve potential international trade effects that should be in the forefront of the minds of developing country policy makers when deciding on genetically-modified crop adoption. Prospective adopters must consider the effects of two types of international trade issues. First, the method of adopting intellectual property that is embodied in genetically-modified crops will affect trade relations with the host country of the intellectual property innovator. Second, the adoption of genetically-modified crop technology may affect international trade relations with consumer nations. Policy-makers are faced with deciding whether the potential benefits of adoption outweigh potential costs.[3]

This essay develops a framework in which to analyse the international trade effects of adopting genetically-modified crop technology. An economic surplus model is constructed that can account for various methods of adoption and a range of export market responses. The model provides a simple and comprehensive method for developing countries to frame the policy decision of adopting genetically-modified crop technology. Every country's decision is different - there exists no absolute ranking of policies. Each country's decision depends on its specific market conditions and probable export market effects.

[2] This essay focuses on the economics of genetically-modified technology, not the science. Chapter 1.2 provides some background on some of the important scientific concerns about genetically-modified technology.

[3] Another consideration is that a developing country that bans GM crop technology may offend the GM innovating country and risk trade action through the World Trade Organisation (WTO). This hypothetical case is not considered in this essay.

The results of this essay's analysis emphasise the importance of three factors. First, a developing country's ability to enforce intellectual property rights is central issue in determining the method of adoption. Second, a levy on genetically modified technology may be a transaction-cost efficient method of avoiding trade actions from innovating countries. Third, prospective genetically-modified crop adopting countries must have an expectation of the response of its export markets to genetically-modified foods. The first and third results are worrisome for developing countries. The institutions that facilitate capacity in each of these factors are lacking in many, if not most, developing countries.

1.2. Developing Countries and Trade in Genetically-Modified Products

This chapter contains two sections. The first section discusses the relevance of genetically-modified (GM) crops to the economic growth of developing countries. This section also includes an overview of the extent of GM crop adoption in developing countries. Section two provides some background on the international trade rules that may be relevant to trade in GM goods (GMGs).

1.2.1. Developing Countries and GMGs

Biotechnology has the potential to provide considerable benefits for developing countries. As with the Green Revolution of the past several decades, the "gene revolution" shows great promise to initiate and speed economic growth and increase welfare in developing countries. There are two primary channels through which GM crop technology can improve welfare. The first is by increasing food security in developing countries by creating crops that are more nutritious, more resistant to weather and pest shocks and less costly to produce (Haggui). There also exists the potential for dynamic benefits as GM technologies are used to initiate Green Revolution-style breeding programs that produce varietal improvements best suited to local agronomic conditions (Evenson).

GM technology may also improve welfare by enhancing what might be an underlying comparative advantage in developing countries. Many developing countries have natural comparative advantages in the production of some agricultural crops because of endowments and low labour costs (Dahlsten). GM crops that increase the sector's productivity deepen developing countries' natural cost advantages and render their exports more competitive. Agricultural exports from developing countries may grow and increase welfare in developing countries.

GM technology can also benefit developing countries by increasing productivity growth where intensive agriculture is practiced. Productivity growth on land that is already under cultivation reduces pressure to expand the quantity of land under cultivation (extensive growth). More food can be grown on the same quantity of land, thus preserving biodiversity (James, 2003). There are also substantial environmental benefits that can arise from the introduction of GM crop technology. Qaim and Zilberman estimate that bacillus thuringiensis (Bt) cotton crops in India require 70% less chemical pesticide than non-GM (NGM) crops in the same region. A reduction in pesticide requirements is particularly beneficial in developing countries where farmers face credit shortfalls that prevent the purchase of chemical pesticides.

The GM technology considered in this essay is of the variety that reduces production costs, not the variety that improves the quality of the food products. Also, the GM adoption decision is considered in an international setting, so that there exists a world food price. For these reasons, the GM technology under consideration in this essay is most likely to increase the adopting country's welfare by enhancing a comparative advantage and increasing exports. Chapter 1.3 explains the manner in which the technological advancement affects production costs.

GM crops are being adopted quickly in developing countries. Developing countries currently account for an estimated 30 percent of global GM hectarage (James, 2003), which is up from below 5 percent just six years ago. It should be noted that the 30 percent estimate is conservative because it includes only a modest approximation of GM soybean crops in Brazil; the true figure is likely higher (James, 2004). Though the level of GM planting in the US (accounting for nearly two-thirds of global hectarage) dwarfs that of all other countries, several developing countries figure prominently in the list of the world's largest GM crop producers. Argentina, Brazil and China account for nearly 20 percent of global GM crop hectarage.

Several developing countries, particularly in Asia, allocate a large sum of public funds to research in GM crop technology. In particular, China, India, Philippines and

Thailand fund public research into technology such as virus-resistant crops (Skerritt). Gray, McNaughton and Stovin estimate that more than 90 percent of GM research is publicly funded in developing countries. Despite this evidence of public research, the vast majority of global GM adoption consists of GM technology that has been created by private firms in the US. The model in this essay illustrates how the developing country's adoption of foreign-owned intellectual property affects welfare.

1.2.2. Trade Rules and Genetically-Modified Goods

There are three sources of potential trade conflict for a country that adopts and trades GMGs. The first is the WTO agreement that governs international protection of intellectual property (IP) rights, the Trade-Related Aspects of Intellectual Property Rights (TRIPS) agreement. The TRIPS agreement can impact the method by which a country decides to adopt GM technology. The second and third sources of potential conflict are the WTO's rules that govern trade in plants and animals, and rules governing labelling requirements. These two agreements, the Agreement of the Application of Sanitary and Phytosanitary Measures (SPS) and the Agreement on Technical Barriers to Trade (TBT) impact the export markets of a country that has adopted GM crop technology. The relevant aspects of these three agreements are discussed below.

The TRIPS agreement is a mandatory component of the WTO; all member countries must be signatories to the agreement. Developed member countries of the WTO pushed hard for the TRIPS agreement to be part of the Uruguay round of trade negotiations as a method of enforcing their domestic IP rights outside their domestic borders. The key to the TRIPS agreement is cross-retaliation with other WTO agreements. If the WTO's Dispute Settlement Body (DSB) determines that a member country is not protecting another member's IP rights (which are enforced for twenty years in the case of patents and fifty years in the case of copyrights), then retaliatory trade measures are allowed under one of the WTO's other agreements. The host country of the IP innovator can impose, on the offending country, trade sanctions that would otherwise be in violation of its General Agreement on Tariffs and Trade (GATT) obligations. Such sanctions would presumably take the form of (often prohibitive) tariffs on imports from the offending nation.

Developed member countries supported cross-retaliation because it was believed to be one of the only methods of enforcing IP rights extraterritorially. Developing countries rarely have IP to protect, so the threat of a developed country retaliating against a pirate industry in a developing country by not protecting developing country IP rights is rarely credible. Developed countries hoped that the imposition of tariffs would coerce developing countries into enforcing foreign IP rights.

The TRIPS agreement is relevant to the GM adoption decision because GM traits that are engineered into GM crops are considered private IP that is owned by innovating firms. Roundup Ready® soybeans that are produced by Monsanto and Bt maize produced by Monsanto, Novartis and Pioneer Hybrid International are patented products whose IP is protected, at least in theory, by the TRIPS agreement. The nature of agricultural crops, however, makes pirating this IP simple from a practical perspective. Farmers or seed breeders can save GM seeds and initiate their own breeding programs that provide a stream of GM seeds in the future.[4] GM seed companies circumvent this problem in developed countries by enforcing contracts with farmers, many of which forbid saving seeds; farmers are legally compelled to repurchase new seeds every year. In developing countries where transaction and monitoring costs are high, IP innovating firms may find it difficult to enforce such contracts[5]. Developed countries therefore hope to use the threat of trade sanctions under the TRIPS to coerce developing countries into enforcing IP rights.[6]

Once an exporting country has adopted GM crop technology, it may face adverse consumer responses from importing nations. Importing nations with consumer concerns about GMGs may seek to restrict imports from GMG-producing countries or impose mandatory labelling requirements on imported GMGs. The WTO's SPS and TBT agreements house the rules that apply to such trade actions.

Before explaining how the SPS and TBT agreements apply to trade in GMGs, it is worth noting that the WTO is ill-equipped to deal with trade disputes about GMGs. The WTO was set up to address protectionism from producers, and as such was initially concerned with the reduction of border measures (Gaisford, et al.). However, it is not

[4] Genetic Use Restricting Technologies, or "terminator genes" that render GM seeds sterile have been abandoned by seed companies in response to public pressure (Wright) and the Consultative Group on International Agricultural Research has officially rejected the use of terminator genes (Pinstrup-Anderson and Cohen).

[5] This point is discussed further in chapter 1.4.

[6] The probable success of cross retaliation as a means of protecting IP rights is discussed in chapter 1.4.

traditional producer protectionism that fuels most calls for import restrictions on GMGs; rather it is consumers and environmentalists that sustain efforts to curb trade in GMGs (Gaisford, et al.). The WTO does not contain a forum for dealing with consumer requests for trade protection, so such requests are handled by WTO constituent agreements that are best (or perhaps the least bad) suited to address the issues presented by GMGs.

The SPS agreement contains rules that pertain to food safety; specifically, rules that attempt to prevent overly strict health and safety regulations from being used as an excuse for domestic protection. It is the SPS agreement that will be asked to decide on the legitimacy of trade actions against imported GMGs that are imposed in the name of food safety. The key to the SPS agreement is that it only allows trade actions that have scientific justification.[7] Trade measures are allowable under the SPS, but such measures must be based on a scientific consensus and must minimise trade disruptions. The scientific consensus must include agreement on the need for a restriction, agreement on the risk associated with not imposing a restriction and agreement on the point at which enough science has been conducted to reach a conclusion on the safety of the GMG in question (Kerr). The requirement of a consensus on "enough science" having been done is troublesome for a rules-based agreement, however (Kerr). A concerned importing nation can always make the claim that not enough science has been done to reach a reliable scientific conclusion. No consensus will exist and trade restrictions on GMGs are justified.

[7] Article 5.7 of the SPS also allows provisional trade measures if there exists insufficient scientific evidence to form a conclusion. Such measures could be consistent with the "precautionary principle".

The TBT agreement seeks to ensure that “regulations, standards, testing and certification procedures do not create unnecessary obstacles” to trade (WTO). In the context of GMGs, the TBT agreement contains rules that govern labelling requirements on food. Like the SPS, the TBT demands that labelling requirements be based on scientific evidence of risk. If a food product poses a safety concern (allergies, for example), then labelling requirements are allowed under the TBT or the SPS. However, it is what the TBT agreement does *not* allow that is likely to lead to disputes between WTO member countries. The TBT agreement does not allow labelling requirements based on consumers’ right to know (Isaac, Phillipson and Kerr). Labelling requirements that are based on consumers’ right to know may violate one of the WTO’s fundamental tenets; “like” products. Specifically, an importing country cannot impose labelling requirements on GM imports just because they have been produced using GM technology. If there is no difference between the final product, regardless of the production method, then the TBT agreement does not permit violation of the “like” products tenet. Such measures are only allowed if it can be demonstrated that the GM production technique produces a different product which presents a risk to the importing country.

The SPS and TBT agreements are intended to provide regulatory predictability to exporters. WTO member countries that fulfill their WTO obligations should not impose unwarranted trade restrictions on GMGs. However, the political reality is that the WTO is a voluntary agreement entered into by member countries and many importing countries appear willing to ignore their obligations in response to domestic political pressure (Kerr and Hobbs). Those countries considering GM adoption must therefore be prepared to encounter trade measures that are not WTO-compliant. The consequences of this are discussed in chapter 1.4.

Another issue complicating the international trade of GMGs is the presence of overlapping jurisdiction between WTO agreements and various multilateral environmental agreements (MEAs). Situations may arise wherein one side of a trade dispute is a member of an MEA that outlines rules for trade in GMGs and the other side is not. If both sides are member countries of the WTO, then it is unclear whether the MEA's or the WTO's rules take precedence. This issue is beyond the scope of this essay - see Isaac and Kerr or Issac, Phillipson and Kerr for more information on such conflicts.

1.3. Model

This section introduces the analytical framework in which GM adoption policies can be analysed. The model provides a framework in which to measure welfare effects of the GM adoption decision in the adopting country. The model is a partial equilibrium economic surplus model and considers only the welfare effects in the industry that adopts GM technology. A basic methodology for analysing welfare effects along a supply chain is presented, and this methodology is then extended to allow for international trade and various adoption policies.

There exist several valid criticisms of economic surplus methods. Alston, et al. provide a comprehensive overview of the primary drawbacks of such methods, and it is worth emphasizing those which are particularly germane to this essay. First, a net welfare gain that is observed in an economic surplus model is no guarantee that everybody is better off. If welfare increases in a comparative static analysis, then the change is Kaldor-Hicks superior; that is, there exists enough new welfare that the winners could compensate the losers so that everybody is better off. The reallocation that is necessary for Pareto superiority may not occur, and presents a particularly large obstacle in developing countries. The institutions that would facilitate such reallocation are "conspicuous by their absence" and transaction costs are very high in developing countries (Hobbs and Kerr).

Another relevant concern is the type of supply shift that is initiated by new technology. Such a shift could be parallel, divergent or convergent. Each of these shifts applies to different types of technological advances and has different implications for changes in welfare. This essay considers a parallel supply shift in the forthcoming model, and the justification for a parallel shift is provided in section 1.3.3.

Despite the shortcomings of economic surplus methods, welfare analysis remains one of the most useful pieces of an economist's toolkit, and is likely the best method for analysis of this kind (Alston, et al.).

This essay does not analyse possible health or environmental aspects of GMGs. Accordingly, the following preliminary assumptions are made: 1) GMGs increase productivity (by reducing costs for the same amount of output) and 2) there are no negative externalities (health or environmental) associated with the use of GM crops. While the first assumption is not overly strong, the second assumption may be contestable. The concerns about the safety of GM products are well documented (see MacFarlane for a discussion). Though GM crops have yet to be proven unsafe for human health or the environment, the technology is new enough that possible negative complications have not yet had time to present themselves. A specific concern about GM crops is that herbicide tolerant and pesticide resistant seeds will enable the evolution of “super weeds” and “super bugs” that are resistant to all herbicides and pesticides.

Food safety is another concern. Diseases that lie dormant for long periods and cancers that take years or decades to develop might not appear for several years after ingesting GM foods. The technology is too new to form a definitive conclusion about these possible effects. There is no evidence linking GM crops to such health risks, however many scientists encourage a precautionary approach.

This analysis does not evaluate the science of these concerns or the rationality of consumer concerns given the scientific evidence. Such concerns are assumed to be reflected in exogenous consumer preferences. Assuming no externalities allows the focus to be placed on the primary economic concern - the potential international trade implications of adopting GM technology.

The forthcoming model develops policy tools that can be used for an adoption decision in most any country, developed or developing. However, a few key assumptions are made that render the policy analysis particularly relevant to developing countries. First, the considered GM technology is assumed to be developed in a foreign country. Since the innovator (most often a biotechnology firm in the US) of the IP is not located in the developing country, the adoption decision involves either pirating the technology or importing and buying the technology from the IP innovator abroad. Any proceeds accruing to the foreign IP innovator must be considered in the welfare analysis. Second, the adopting nation is modelled as a small country, so that its production and exports do not affect the world price.[8]

The analysis is conducted from the adopting-country's perspective. This country is referred to as DC (a generic method of referring to "developing country"), and it is DC's economic welfare that forms the basis of policy decisions. The forthcoming analysis could be applied to any number of different crops, but reference is made to adoption of GM soybeans, for the sake of convenience. DC produces and exports NGM soybeans, and the policy question surrounds the adoption of GM soybean seeds. The example of soybeans is used because of its relevance to current developing country situations. Several South American nations have recently approved domestic use of

[8] The case where the adopting developing country affects the world price for GM products is considered in a later section.

GM soybeans and other nations are in the decision stage of approving GM technology (James, 2004). Rice is another topical example. China is currently in the stages of developing and approving GM rice for domestic production (James, 2004). The analysis is presented in an order that follows the soybean supply chain in chronological order - that is, the welfare effects of importing/buying GM seeds and growing GM soybeans are derived first, followed by the welfare effects of exporting GM soybeans. However DC's policy analysis should begin at the other end of the supply chain, with DC determining the probable effects on its export market before moving to the decision of whether (and in what manner) to adopt GM technology. This point is addressed further in chapter 1.4.

1.3.1. A Simple Fixed Proportions Model

The policy decision under consideration in this essay affects several related input markets. The production and exportation of GM soybeans involves seeds, land, chemicals, elevator and transport services. The following analysis presents the policy decision in a multi-input, single-output fixed proportions model. The primary benefit of a fixed-proportions model is that it allows a close analysis of how policies and technological factors in one stage of the supply chain affect other stages through vertical linkages.

We assume that three inputs are required in the production and exportation of beans. Seeds (S) are the primary input and are planted on farm land (L). The third input is an aggregate input that includes the cost of herbicides and other inputs per hectare of land

(C). The other input costs per hectare include elevator and transportation services. These inputs are used in a fixed proportion to produce an equivalent quantity of beans (B). That is, $S + L + C = B$. This states that, for example, a fixed quantity of seeds (measured in tonnes), one hectare of land and one hectare's worth of other inputs are required to produce one tonne of beans. The fixed-proportions framework does not allow for input substitutability - inputs must be applied in the same quantity to produce a fixed amount of beans. This assumption is reasonable for the purposes of this model. The GM technology considered in the forthcoming analysis reduces the marginal cost of certain inputs, but does not change the quantity of inputs required to produce a fixed level of output.

The units of each input can be adjusted so that $S = L = C = B$. Scaling factors that could be attached to each input (so that, for example, 0.001 tonnes of seed are required to produce one tonne of beans) are omitted for notational convenience. The forthcoming analytical welfare derivations are not affected by this omission, but if empirical estimates of demand and supply parameters were available, then scaling factors would be required to calculate cardinal welfare values. The markets for all inputs are assumed to be perfectly competitive[9] so that the supply price of the final good (beans) is equal to the sum of the costs of each input. That is, $P_B = P_S + P_L + P_C$. The supply curve for beans can then be derived as the vertical summation of all input supply curves. Consider the following inverse supply curves for seed, land and other input costs per hectare of land, respectively.

---

[9] This restriction is relaxed in a later part of the analysis.

$$P_S = \mu \tag{1.1}$$

$$P_L = \gamma + \lambda L \tag{1.2}$$

$$P_C = \chi \tag{1.3}$$

The inverse supply curves in equations (1.1) to (1.3) represent aggregate supply curves for each input in DC. Seeds are assumed to be supplied at a constant marginal cost of μ, land supply is upward sloping, representing it as a Ricardian input (fixed in quantity), and the marginal cost of other inputs per hectare of planted land is assumed to be constant across output levels. Adding equations (1.1), (1.2) and (1.3) yields DC's inverse bean supply curve:

$$P_B = \chi + \gamma + \mu + \lambda L\,, \tag{1.4}$$

and because input units can be adjusted so that $L = B$, we have

$$P_B = (\chi + \gamma + \mu) + \lambda B\,. \tag{1.5}$$

Supply of beans in DC is determined according to equation (1.5). A bean demand function is now required to establish equilibrium in these related markets[10]. Consider DC's inverse bean demand as

[10] Price is determined locally in the introductory model, indicating a closed economy. The model introduces international trade in the next section.

$$P_B = \omega - \phi B. \tag{1.6}$$

The DC bean market is in equilibrium where bean supply is equal to bean demand:

$$(\chi + \gamma + \mu) + \lambda B = \omega - \phi B, \tag{1.7}$$

or

$$B^* = \frac{\omega - \chi - \gamma - \mu}{\lambda + \phi}. \tag{1.8}$$

DC equilibrium bean output equals equation (1.8), which is a function of the slope and intercept parameters of the related industries' supply functions and of the bean demand function. The equilibrium price in this market is solved as

$$P_B^* = \omega - \phi\left(\frac{\omega - \chi - \gamma - \mu}{\lambda + \phi}\right). \tag{1.9}$$

Derived demand curves for each input market can now be found using the equilibrium bean price and the knowledge that each input is supplied competitively. Producers are willing to pay for each input unit an amount equal to the final price of beans minus the price he must pay for each of the other two inputs required in production. For example the derived demand for seed inputs is equal to the bean price minus the supply price of other inputs per hectare minus the supply price of land. As such, inverse derived demand functions are equal to the vertical difference between the equilibrium price for beans and the supply functions of the two other inputs.

$$P_S = \omega - \phi\left(\frac{\omega - \chi - \gamma - \mu}{\lambda + \phi}\right) - \gamma - \lambda L - \chi \,. \tag{1.10}$$

Equation (1.10) represents the inverse demand function for seeds. Since $L = S$ in the fixed proportions model, and substituting equation (1.8) into equation (1.10), the inverse derived demand function for seeds is

$$P_S = \left(\omega - \phi B^* - \gamma - \chi\right) - \lambda S \,. \tag{1.11}$$

Inverse derived demand functions for land and for other inputs per hectare can be constructed in a similar fashion:

$$P_L = \left(\omega - \phi B^* - \mu - \chi\right) \tag{1.12}$$

$$P_C = \left(\omega - \phi B^* - \mu - \chi\right) - \lambda C \,. \tag{1.13}$$

Equilibrium prices in each market can be determined by recalling that, after adjusting units, $S = L = C = B$ so that $S^* = L^* = C^* = B^*$. Equilibrium output levels in each market can be inserted into the relevant inverse demand or supply curves to obtain equilibrium prices in each market.

The related markets that are outlined in equations (1.1) through (1.13) can be illustrated graphically. Figure 1.1 presents the markets for DC's beans and bean inputs. Each panel represents a different stage in the vertically linked supply chain for beans. The exogenous portions of the model are seed supply, land supply, other inputs per hectare supply and bean demand. The bean supply function and demand functions for all inputs are derived using the aforementioned methodology. The equilibrium price for beans is equal to the sum of the equilibrium cost for seeds, land and other inputs per hectare.

Though each market is competitive, there exist rents to each segment in the form of consumer and producer surpluses. However, one of the chief benefits of a fixed proportions model is that the entire industry's rents can be derived from analysis of the end market. Since the supply and demand curves in the bean market represent the totals of each supply-chain segment below, consumer and producer surplus measures taken from the bean market provide a measure of total welfare[11]. Total welfare in the DC bean industry is measured by the sum of consumer surplus (area abP_B^W) and producer surplus (area $P_B^W bc$) in panel 1 of Figure 1.1.

1.3.2. International Trade in a Fixed-Proportions Model

The simple model outlined above can be adapted to allow for international trade by introducing world bean supply and demand functions. Only the world final-product market functions (rather than all constituent input supply and derived input demand

[11] Benefits to each segment of the supply chain can be deduced from welfare measures in each panel of Figure 1.1.

curves) are presented for reasons of brevity. World bean inverse demand is represented as

$$P_B = \Omega - \alpha B \tag{1.14}$$

and world bean inverse supply as

$$P_B = \theta + \beta B\,. \tag{1.15}$$

Equation (1.14) reflects global demand for beans, and is necessarily flatter than DC's demand for beans; that is, $\alpha < \phi$. Global bean supply is the horizontal sum of all nations' bean supply curves and is necessarily larger than DC bean supply, so that $\beta < \lambda$. A preliminary assumption is made that the intercept of the world bean inverse supply curve is equal to the intercept of DC's inverse supply curve. That is, $\theta = \chi + \gamma + \mu$. Setting world supply equal to world demand yields the world bean price:

$$P_B^W = \Omega - \frac{\alpha(\Omega - \theta)}{(\alpha + \beta)} \tag{1.16}$$

and world bean consumption

$$B^{W*} = \frac{(\Omega - \theta)}{(\alpha + \beta)}\,. \tag{1.17}$$

The model assumes that only the final product in the bean industry is traded internationally. Inputs, including other services (pesticide, fertiliser, elevator and transportation), land and seeds, are supplied and used domestically so that prices are determined locally in the DC market[12].

The world bean price is used to derive the input demand curves for DC's bean industry instead of equation (1.9). Subtracting the relevant supply functions in equations (1.1) through (1.3) from the world bean price in equation (1.16) yields derived inverse demand functions for seeds, land and other inputs per hectare, respectively:

$$P_S^{DC} = \left[\theta - \frac{\alpha(\Omega - \theta)}{(\alpha + \beta)}\right] - \lambda S \tag{1.18}$$

$$P_L^{DC} = \Omega - \frac{\alpha(\Omega - \theta)}{(\alpha + \beta)} - \chi - \mu \tag{1.19}$$

$$P_C^{DC} = \left[\Omega - \gamma - \mu - \frac{\alpha(\Omega - \theta)}{(\alpha + \beta)}\right] - \lambda C\,. \tag{1.20}$$

At a price of beans given by equation (1.16), DC produces beans according to its supply function given in equation (1.5). DC bean output is

[12] The case where GM seeds are imported is considered at a later stage.

$$B^{DC*} = \frac{\Omega - \frac{\alpha(\Omega - \theta)}{(\alpha + \beta)} - \chi - \lambda - \mu}{\lambda} \\ = \frac{P_B^W - \theta}{\lambda}. \qquad (1.21)$$

The fixed proportions model in Figure 1.1 is adapted to allow for international trade in Figure 1.2. S_B^W represents global bean supply, and is necessarily flatter than DC bean supply because world supply is equal to the horizontal summation of all constituent nations' supply. Global bean supply intersects global bean demand, D_B^W, to establish the equilibrium price of equation (1.16). DC's derived input demand curves are determined by subtracting the relevant input supply functions from the world bean price.

DC's domestic welfare can be measured by the sum of consumer surplus (area acP_B^W) and producer surplus (area $P_B^W bd$) in the top panel of figure 1.2. These two areas account for all rents accruing down the vertically linked supply chain in DC's bean industry.

1.3.3. Technological Change in a Fixed Proportions Model

We now consider the effects of a technological advancement in the form of a GM seed trait. There are several types of genetic modifications that would be considered a technological improvement from an economist's perspective. GM traits can produce foods with longer shelf lives, higher nutritional value or appealing aesthetic characteristics. Such traits are product-quality based and affect market demand, not supply. This essay, however, focuses on GM traits that impact bean supply conditions. Specifically, GM traits that reduce input costs along the bean supply chain are considered.

Two such traits are most prevalent. The use of a herbicide tolerant gene in soy, maize, cotton and canola crops has created Roundup Ready® crops that are tolerant to post-emergence application of glysophate herbicides. Resistance to glysophate allows farmers to reduce their herbicide costs by applying one broad-spectrum chemical to their crops. Application of a broad-spectrum herbicide also reduces tillage requirements. The result is decreased production costs per hectare of planted crop. The second most common GM characteristic in agriculture is the Bt trait that attributes modified crops with deterrence against insect pests. The Bt trait is commonly found in maize and cotton crops, and reduces production costs per hectare by decreasing reliance on chemical pesticides.

This essay focuses on cost-reducing GM traits for two reasons. The first is that such traits account for the vast majority of global GM hectarage. Herbicide tolerant crops account for 72 percent of global GM crops and Bt for 19 percent (James, 2004). The second reason is that the adoption of herbicide tolerant and insect resistant crops are germane to current policy decisions in developing countries. Several South American nations (including Brazil, Argentina and Paraguay) have recently adopted and approved GM crops, or are considering doing so in the near future. China is likely to approve and adopt Bt rice over the next year (James, 2004). The potential welfare effects of such policy decisions should be of great interest to decision-making authorities in these countries.

To analyse the effects of GM adoption in DC, we assume that farmers are faced with the initial option to adopt GM seeds. Farmers decide on the quantity of GM seeds to purchase according to their derived GM seed demand curve, and allocate an equivalent amount of land and other services per hectare to the production of GM beans. This decision is based on the price received for GM beans and on the input costs along the GM bean supply chain. The bean price received and the incurred production costs change in the forthcoming scenarios and results in different adoption decisions.

The introduction of a GM herbicide tolerant or insect-resistant trait into bean seed reduces the quantity of chemical herbicides or pesticides that are required on a given hectare to achieve an equivalent yield. Put another way, the GM trait lowers the cost of other inputs per hectare of planted land. Such a trait can be modelled as a parallel downward shift of the inverse supply curve for other services per hectare[13]. The marginal cost of other inputs is assumed to be constant across output levels. The GM trait maintains constant marginal costs, but at a lower level. Equation (1.3) then becomes

[13] Two initial assumptions are made at this stage. First, producers do not pay a royalty or monopoly price (a technology use fee) for GM seeds. The effects of royalties and higher GM seed prices are taken up in section 1.3.4. Second, there is no adverse consumer reaction to GM beans. The effects of various consumer responses are examined in section 1.3.5.

$$P_C = \varepsilon\chi, 0 < \varepsilon < 1, \tag{1.22}$$

where ε represents the downward shift of the inverse supply curve for other inputs per hectare. Equation (1.22) is now used to generate the inverse bean supply curve, which can be represented as

$$P_B = (\varepsilon\chi + \gamma + \mu) + \lambda B\,. \tag{1.23}$$

Notice that the inverse supply curve in equation (1.23) is necessarily lower, or further out, than the pre-GM inverse supply curve in equation (1.5) because $\varepsilon < 1$. DC can now produce beans at a lower cost than in the pre-GM situation.[14] It is assumed that GM seeds can be supplied at the same marginal cost as NMG beans, so that equation (1.1) remains relevant. Likewise, the marginal cost of land remains the same as in equation (1.2).

DC is considered a small country in the global economy, so its output of beans does not affect world supply conditions, and therefore does not change the world price. We therefore determine DC's GM bean output using equation (1.16) according to the inverse supply function in equation (1.23). DC's equilibrium GM bean output is

[14] From this point on, we will consider the international framework of equations (1.1) through (1.3), (1.5),(1.6) and (1.14) though (1.21) as the pre-GM baseline. This baseline is used for comparing welfare effects of policy decisions regarding GM adoption and possible trade effects.

$$B_{GM}^{DC*} = \frac{\Omega - \frac{\alpha(\Omega - \theta)}{(\alpha + \beta)} - \varepsilon\chi - \gamma - \mu}{\lambda}$$
$$= \frac{P_B^W - \varepsilon\chi - \gamma - \mu}{\lambda} \tag{1.24}$$

which is necessarily larger than pre-GM bean output because $\varepsilon < 1$.

It is intuitive that DC completely adopts GM seeds and there is no NGM bean production in this scenario. The world price is the same for both products and the cost of producing GM beans is lower; there is no incentive to produce NGM beans.

Figure 1.3 illustrates a graphical representation of GM seed adoption as described above. The introduction of GM seeds reduces the marginal cost of other inputs per hectare, pushing S_C down to $S_{C,GM}$. The supply of beans in the GM market (the vertical sum of all input supply curves) then shifts out to $S_{B,GM}^{DC}$. DC now produces more beans (all of them GM) than in the pre-GM situation. DC's NGM bean production falls to zero.

The change in DC's domestic welfare that results from GM adoption can be measured as the difference between total welfare in the top panel of figure 1.3 and total welfare in the top panel of figure 1.2. Consumer surplus remains unchanged after the introduction of GM technology - DC consumes the same quantity of beans at the same world price.[15]

[15] DC consumers are assumed to be indifferent between GM and NGM beans.

Producer surplus unambiguously rises, however, since DC produces and exports more beans. DC unambiguously gains from the adoption of GM seed technology in this case.

This situation represents the best-case scenario for DC. DC benefits from the cost-reducing technology without facing trade actions on either end of its bean supply chain (purchasing seeds or exporting beans). GM beans can be exported alongside NGM beans at the same price to the same foreign consumers. This result is dependent on there being no consumer preference for NGM beans over GM beans and on none of DC's export markets imposing trade restrictions or labelling requirements. The benefits also depend on DC pirating the GM technology from the foreign innovator and producing GM seeds for the same marginal cost as is required to produce NGM seeds. Subsequent cases examine the effects on DC welfare of actions by the innovating firm and of export restrictions.

Now that the basics of the fixed proportions model have been established to accommodate international trade and technological change, we turn to modeling policy options that determine the welfare implications of DC's decision to adopt GM seeds.

1.3.4. GM Adoption Scenarios

We now consider three scenarios under which GM technology can be adopted in DC. The policy options include pirating GM technology, paying the IP innovator the monopoly price, and negotiating a levy on production of beans while maintaining a black market for GM seeds. Each option is described in detail below and the welfare gains from each option are derived and compared. If empirical estimates of the relevant supply and demand functions were available, then welfare values could be calculated to assist in policy decisions.

Case 1 - Pirate GM Seeds

Case 1 considers the scenario outlined above in which DC adopts GM technology, faces no constraints on its export market and the marginal cost of seeds does not rise above the marginal cost of NGM seeds. The results are identical to the baseline case of technological change in section 1.3.3. Such a scenario may seem unrealistic and "too good to be true". However, it is just such a situation in which countries such as Brazil have found themselves in recent years. Brazil had not yet authorised the use of GM soybeans, and therefore maintained unfettered access to EU markets. However James (2004) estimates that almost one-quarter of Brazil's soybean crop is GM - an estimate that James acknowledges is almost certainly low. Furthermore, Brazil's GM soybean crops are not grown from seeds purchased from Monsanto, the patent-holder for herbicide tolerant soybean seeds. Rather, Brazilian seeds are purchased on the local market from black market dealers who are believed to have smuggled GM soybean seeds from Argentina a decade ago (The Western Producer). Brazilian farmers need not pay a higher price for the technologically advanced seeds - a situation that is considered at a later stage of this analysis.

DC's welfare can be deduced from the top panel of figure 1.3. Consumer welfare remains equal to area acP_B^W and producer surplus grows to area $P_B^W bd$. There is no economic surplus in the NGM market since production has fallen to zero.[16] Welfare in the post-GM scenario is higher than in the pre-GM scenario because producer surplus is unambiguously larger.

The result that GM adoption is complete warrants some discussion. GM adoption is estimated to be incomplete in several developing countries, including Brazil (James, 2004). The model's result of complete GM adoption is, however, a function of the comparative static nature of the welfare model. The model illustrates two equilibriums, pre-GM and post-GM. There is no accounting for a lag in adoption that would explain an intermediate stage wherein adoption is not complete. Argentina, where GM soybeans were adopted earlier than in Brazil, is estimated to be very near 100 percent GM and Brazil's GM area is growing quickly (James, 2004). This lends credence to the suggestion that the adoption process is dynamic and does not occur instantaneously. A lag in adoption could be due to a several factors including shortage of seeds, slow transmission of information to farmers or precautionary decision making by farmers (farmers do not adopt until they evaluate the success of neighbouring farms that have adopted). The key result from the model is that there exist economic incentives for farmers to adopt GM seeds. Comparative statics suggest that the new equilibrium (complete GM adoption) is attained once obstacles are overcome.

Case 2 - DC Pays the Monopoly Price for GM Seeds

[16] There is an underlying assumption that all farms are equally agronomically suited to GM seeds. The GM trait under consideration in this essay is assumed to reduce production costs for all farms by the same amount - a reduction in herbicide or pesticide cost per hectare.

If the GM innovator can successfully protect his IP rights in DC, then DC's farmers are forced to purchase GM seeds directly from the innovating firm (or one of its agents). Presuming that the GM technology is patented and that the innovator is the only firm selling such seeds, the innovator holds monopoly power over the GM seed market and sells seeds at a price and quantity so as to maximise profits. The monopolist's profit-maxinising price depends, however, on the level of cost savings that the GM technology introduces to the supply chain. That is, the monopolist's ability to earn profits by constraining seed output and raising seed price depends on the amount that the supply curve for other inputs per hectare shifts down (i.e. the size of ε). If the innovation reduces costs by a large enough amount to be considered *drastic*, then the monopolist can capture the entire seed market by selling a larger quantity of GM seeds at a lower price than in the pre-GM equilibrium (Moschini and Lapan). The price at which the downward sloping portion of the monopolist's marginal revenue is equal to marginal cost, once adjusted for the increased efficiency of the new technology, is below the pre-innovation price for NGM seeds. The monopolist therefore faces no effective competition from NGM seeds and captures the entire market.

If GM technology does not reduce the cost of other inputs per hectare enough to allow the monopolist to charge a price where marginal cost equals the downward sloping portion of marginal revenue, then the innovation is *nondrastic* (Moschini and Lapan). In the case of a nondrastic innovation, the monopolist must consider competition from the existing NGM technology when setting price. No farmer will adopt the GM technology as long as the price for the GM seeds is higher than the price of the NGM seeds *plus* the cost saving arising from adopting the new technology (i.e. the efficiency adjustment). The world price for both GM and NGM beans is the same so there is no economic incentive for farmers to adopt GM seeds unless it affords them a cost saving. The monopolist's ability to constrain output and increase price relative to the competitive equilibrium is therefore constrained in the case of a nondrastic innovation. The profit-maximising solution is no longer to set price where marginal cost is equal to the downward sloping portion of marginal revenue, because no farmer will pay that price for GM seeds; NGM seeds sell for less (adjusted for efficiency). The monopolist must therefore charge a price below the (efficiency-adjusted) competitive NGM seed price to sell any seeds. This becomes the monopolist's profit-maximising strategy in the case of a nondrastic innovation and results in his capturing the entire market. An analytical and graphical explanation the monopolist's pricing decision follows.

The world bean price remains as in equation (1.16), and the derived demand for GM beans is calculated by subtracting the supply functions for land and for other inputs from the world bean price. This generates

$$P_S = \left[P_B^W - \varepsilon\chi - \gamma\right] - \lambda S \ . \tag{1.25}$$

Note, however, that equation (1.25) represents demand for GM seeds only beyond the quantity sold in the pre-GM equilibrium. Farmers can acquire NGM seeds for a price of μ after the GM innovation is introduced and will therefore not pay more than μ plus the cost saving introduced by the trait for GM seeds. Therefore, the demand for GM seeds is perfectly elastic at a price of μ up to the pre-GM equilibrium output, as given in equation (1.21). GM bean demand is discontinuous at that level of output and jumps to the level given by equation (1.25) thereafter. The complete derived seed demand curve is then

$$P_S = \begin{cases} \mu \; for \; S \leq \dfrac{P_B^W - \theta}{\lambda} \\ \left[P_B^W - \varepsilon\chi - \gamma\right] - \lambda S \; for \; S > \dfrac{P_B^W - \theta}{\lambda} \end{cases} .^{17} \tag{1.26}$$

Equation (1.26) states that demand for GM seeds is perfectly elastic at μ at output levels below the pre-GM competitive equilibrium, and downward sloping thereafter.

The marginal revenue function facing the monopolist is also discontinuous and is represented as

$$MR_S = \begin{cases} \mu \; for \; S \leq \dfrac{P_B^W - \theta}{\lambda} \\ \left[P_B^W - \varepsilon\chi - \gamma\right] - 2\lambda S \; for \; S > \dfrac{P_B^W - \theta}{\lambda} \end{cases} . \tag{1.27}$$

[17] Note that the demand function remains flat at the pre-GM level of output. If there exist no adjustment costs in switching from NGM to GM seeds, then the NGM seed price is exactly equal to the efficiency-adjusted GM seed price at this level of output; farmers would be indifferent between NGM and GM seeds. That the downward sloping portion of the GM seed demand curve begins after the pre-GM level of output is therefore dependent on there being no incentive for farmers to switch to GM seeds when the efficiency-adjusted seed prices are identical; a small adjustment cost requires that the efficiency-adjusted price for GM seeds must fall marginally for farmers to switch to GM seeds.

To acquire an analytical solution for the monopolist's price and output combination, we initially ignore the prospect of competition from NGM seeds. The innovating firm maximises profits by choosing quantity where marginal revenue (recall that competition from NGM seeds is initially ignored, so that we consider the downward sloping portion of the marginal revenue function of equation (1.27)) is equal to marginal cost. The innovating firm's marginal cost of producing GM seeds is assumed to remain constant at μ. Setting marginal cost equal to marginal revenue and solving for seed output yields

$$S_{GM}^{DC*} = \frac{P_B^W - \varepsilon\chi - \gamma - \mu}{2\lambda}. \tag{1.28}$$

At this level of seed output, the monopolist charges a price of

$$P_S = \frac{P_B^W - \varepsilon\chi - \gamma + \mu}{2}. \tag{1.29}$$

The price in equation (1.29) is determined using the unconstrained, downward sloping demand function from equation (1.25). As such, the price and quantity combinations of equations (1.28) and (1.29) are relevant only in the case of a drastic innovation, in which competition from NGM seeds is not relevant. The GM innovation introduces such a large cost saving that the monopolist is unconstrained in his pricing decision and can set price where marginal cost intersects the downward sloping section of the marginal revenue curve. For this to be the case, the price in equation (1.29) must be below the pre-innovation NGM seed price, adjusted for efficiency; otherwise farmers have no incentive to purchase GM seeds. This requirement can be shown analytically as

$$\underbrace{P_S}_{\substack{monopolist's\\ unconstrained\\ profit-\max imi\sin g\\ price}} < \underbrace{\mu}_{\substack{competitive\\ NGM\\ seed\\ price}} + \underbrace{\chi - \varepsilon\chi}_{\substack{\cos t\\ saving\\ from\\ GM-techno\log y}} . \tag{1.30}$$

Equation (1.30) states that for an innovation to be drastic, the monopolist's profit-maximising price must be less than the pre-innovation NGM seed price plus the cost saving that GM technology introduces into the bean supply chain. Substituting equation (1.29) into (1.30) and solving for ε generates an inequality restriction on the amount that GM technology must reduce production costs in order to be a drastic innovation. If ε satisfies

$$\varepsilon < \frac{\mu + 2\chi - P_B^W + \gamma}{\chi} \tag{1.31}$$

then the innovation is drastic and the monopolist captures the entire seed market.

A drastic innovation results in larger DC seed, and therefore bean, output and a lower DC seed price compared to the pre-innovation situation. Bean output can be solved as

$$B_{GM}^{DC*} = \frac{P_B^W - \varepsilon\chi - \gamma - \mu}{2\lambda} \tag{1.32}$$

through the fixed proportions characteristic of the model; equation (1.32) is larger than the pre-GM output of equation (1.21) if inequality (1.31) is satisfied. Also, the seed price of equation (1.29) is lower than the efficiency-adjusted pre-innovation seed price of $\mu + \chi - \varepsilon\chi$ if the innovation is drastic. Note that GM beans sell for the same world price as NGM beans, and farmers pay a lower price for seeds than in the pre-innovation equilibrium. This results in rents accruing to owners of land and other services as revenue from bean sales is apportioned along the supply chain.

The case of a drastic innovation is illustrated in figure 1.4. Note that only the GM market is illustrated in figure 1.4; the analytical solution shows that GM adoption is complete, so that NGM bean production falls to zero. The technological shift is represented by the supply curve for other inputs shifting down to $S_{C,GM}$, which is below supply curve from the NGM market (drawn as S_C). Demand for land that is used to grow GM crops correspondingly shifts up to $D_{L,GM}$. Demand for GM seeds and marginal revenue are discontinuous in the lower panel as described above. The drastic innovation results in derived demand for seeds shifting out far enough that the monopolist can charge price where marginal cost is equal to the downward sloping portion of the marginal revenue curve. Seed, and therefore bean, output is above the pre-innovation equilibrium but below the pirate equilibrium output of equation (1.24). Note that the demand price of inputs is above the marginal cost of supplying inputs in the middle two panels of the GM market (land and other inputs). This results in economic profits to owners of land and other services that are used in the production of

GM beans. DC's welfare is deduced from the top panel of figure 1.4. Consumer surplus remains unchanged, equal to area abP_B^W and producer surplus rises above the pre-GM level to $P_B^W cde$. The foreign IP innovator's revenue must, however, be deducted from DC welfare in an amount equal to $P_S^M\ fg0$.

Before analysing a nondrastic innovation, we first consider the trivial case in which GM technology does not introduce any cost saving into the bean supply chain.[18] If GM seeds introduce no (negative) cost savings into the bean supply chain, then derived demand for GM bean seed coincides with (lies below) derived demand for NGM seeds. That is,

$$\left[P_B^W - \varepsilon\chi - \gamma\right] - \lambda S \leq \left[P_B^W - \chi - \gamma\right] - \lambda S. \tag{1.33}$$

The inequality in (1.33) can be rearranged to show that if $\varepsilon \geq 1$ then demand for GM seeds is always at or below demand for NGM seeds and farmers have no incentive to adopt GM technology. Recall, however, that ε is restricted to be less than one in equation (1.22); otherwise the GM technology is not a cost-reducing innovation. As such, the trivial case is of interest only because it provides the upper bound (i.e. less than one) on the cost saving introduced by GM technology. We know that if inequality (1.31) is satisfied, then the innovation is drastic, GM adoption is complete, bean output

[18] The trivial case is presented only to introduce boundary conditions on the technical innovation (i.e. the size of ε).

is equal to equation (1.32) in DC and the monopolist sells seeds for the price in equation (1.29).

If ε falls in the intermediate range (less than one but violating (1.31)) then the monopolist must consider competition from NGM seeds when setting price; the price of equation (1.29) is no longer the monopolist's profit-maximising solution. In the case of a nondrastic innovation, the price of equation (1.29) does not satisfy the efficiency adjustment of equation (1.30); price (1.29) is higher than the competitive NGM seed price, adjusted for efficiency. Farmers have no incentive to purchase GM seeds.

The GM innovator must therefore reduce the price of his seeds to induce farmers to adopt GM technology. The monopolist can, however, still capture the entire seed market and earn profits. Recall that as long as ε is less than one, demand for GM seeds lies above demand for NGM seeds and farmers are willing to pay more for GM seeds. The monopolist can therefore capture the entire seed market by charging a price marginally below the competitive price plus the cost savings from GM technology. This can be shown analytically by noting that farmers have no incentive to switch to GM seeds at a price of $\mu + \chi(1-\varepsilon)$, but if the monopolist charges marginally below the efficiency-adjusted equivalent price for NGM seeds then all farmers adopt GM technology. A profit-maximising monopolist innovator of a nondrastic technology therefore charges a price of

$$P_S = \mu + \chi(1-\varepsilon) - e \tag{1.34}$$

where e is a marginally small constant. The price of GM seeds is above the constant marginal cost of producing GM seeds, so the monopolist earns profits. GM seeds are sold according to the downward sloping portion of the demand curve in equation (1.26) in an amount marginally above the pre-GM equilibrium, equal to

$$S_{GM}^{DC*} = \frac{P_B^W - \varepsilon\chi - \gamma - \mu}{2\lambda} + E \tag{1.35}$$

where E is an marginally small constant. The price in (1.34) is below the price of equation (1.29), but the monopolist does earn profits in the case of a nondrastic innovation. As long as the downward sloping portion of derived seed demand is above marginal cost, the average price that the monopolist receives is above the average production cost. Note, however that the monopolist has incentive to undercut the competitive pre-GM seed price by just a marginal amount because the more price falls below equation (1.29) the larger is the difference between marginal cost and marginal revenue. The equilibrium result in the case of a nondrastic innovation is that GM seeds are completely adopted by farmers at the price in equation (1.34) and output marginally above pre-GM output, as indicated in (1.35).

The nondrastic innovation equilibrium is illustrated in figure 1.5. The downward sloping portion of the monopolist's marginal revenue curve does not intersect marginal cost to the right of the pre-innovation level of output. The monopolist therefore charges a price marginally below $\mu + \chi(1-\varepsilon)$ and sells a quantity of seeds marginally above the pre-innovation level of output. As in the case of a drastic innovation, GM adoption is complete and NGM bean production falls to zero in DC. Total DC bean production is just marginally above pre-innovation NGM bean output. DC's welfare is measured as consumer surplus of area abP_B^W and producer surplus of area $P_B^W cde$. As in the case of a drastic innovation, the IP innovator's revenue of $P_S^M fg0$ must be deducted from DC's welfare.

The monopoly scenario of case 2 arises only in situations when the innovator's IP rights can be enforced through an effective monitoring and judicial system. Ineffective IP enforcement is likely to result in a black market and lower prices for GM seeds. Producers will have no incentive to purchase higher-priced GM seeds from the IP innovator and the bean market will revert to case 1. The ability of developing countries to maintain effective IP enforcement mechanisms is questionable, and is discussed further in chapter 1.4.

<u>Case 3 - DC purchases black market seeds and pays a levy to the IP innovator</u>

When the conditions of the producing-country are not amenable to effective IP enforcement, then the IP innovator is faced with a dilemma. The innovating firm can attempt to charge the monopoly price for its product, but will likely be unsuccessful and forced from the market by low-priced black market seeds. This option is unappealing since the innovating firm incurs the costs of marketing seeds and implementing an IP enforcement mechanism, but likely generates no revenue over the long term. A second option is for the IP innovator to do nothing, so that no costs are incurred and DC is left to its black market equilibrium.

A third option, one that has recently been adopted by Monsanto in its dealings with several South American countries, is to negotiate a levy on all beans produced in DC. Such a levy would have to be negotiated with the producing-country government, and enforced at one stage of the bean supply chain. Brazil and Paraguay have recently agreed to a royalty agreement with Monsanto that imposes a levy on all GM products grown in their countries (The Western Producer). Brazil's agreement provides Monsanto with one percent of sales earned from the 2004-2005 crop and two percent of sales from the 2005-2006 crop. Paraguayan farmers have agreed to pay a fixed levy per sack of soy seeds to Monsanto. Brazil and Paraguay's seed markets remain primarily black, but Monsanto now receives a share of proceeds from the use of its IP. Other possible options include applying a check-off levy at elevators where beans are delivered post-harvest or enforcing an export levy on beans as they leave the producing country.

An elevator levy may be the most comprehensive method for an IP innovator to collect royalties on GM technology. If farmers save and replant a portion of their seeds, then a levy on seeds would apply only to those farmers purchasing new seeds. Likewise, an export levy would only apply to beans destined for foreign markets. A levy at elevators would apply to all producers who use elevator services.

The effects of a levy can be incorporated into the fixed proportions model of case 1. We begin from a starting point of pirated GM seeds, as in case 1. A levy is imposed on bean producers in the form of a check-off at the elevator. The levy is a fixed amount per tonne of bean, and therefore increases the marginal cost of other services per hectare by a fixed amount at all levels of output. The inverse supply function of other services is then

$$P_C = \varepsilon\chi + \sigma, 0 < \varepsilon < 1, \sigma > 0 \tag{1.36}$$

where σ represents the per tonne levy. The new DC inverse supply function of beans is

$$P_B = (\varepsilon\chi + \sigma + \gamma + \mu) + \lambda B \tag{1.37}$$

and DC produces beans in the amount of

$$B_{GM}^{DC*} = \frac{P_B^W - \varepsilon\chi - \sigma - \gamma - \mu}{\lambda} \tag{1.38}$$

at the fixed world price.

Bean output under the levy is necessarily less than in the black market situation of case 1 because $\sigma > 0$.

For there to be any NGM bean production in case 3, the levy on GM bean producers must be large enough to more than negate the cost-reducing benefits of the GM technology. That is, $\sigma > \chi(1-\varepsilon)$ must hold. If that is the case, then there is no GM bean production in DC, only NGM production at the original pre-GM level.

The size of the levy (σ) depends on several factors, but is ultimately decided in negotiations between the IP innovator and DC's government. The levy must be large enough to avert trade action by the IP innovator and small enough that DC perceives a benefit in its implementation. Traxler notes that only a fraction of an imposed levy is likely to accrue to the IP innovator; proceeds are likely to be split into at least four portions. The negotiating government is likely to command a share of the levy, in part to offset the costs of negotiating and managing the levy. Also, the stage of the supply chain that collects the levy will need to be compensated for the costs of collection. For example elevator operators who collect a check-off are likely to receive a share of a levy. Finally, the IP innovator may dedicate some portion of the proceeds to agricultural research and development in the adopting country (Traxler).

Case 3 is represented graphically in figure 1.6. Using case 1 as a starting point, DC utilises black market GM seeds and produces only GM beans. The negotiated levy with the IP innovator is collected at the elevator, which is a component of other services per hectare. The levy increases the marginal cost of other services, pushing up the inverse supply curve of other services per hectare from $S_{C,GM}$ to $S_{C,GM}'$ (note that $S_{C,GM}'$

remains below S_C). The inverse bean supply curve shifts up accordingly and the demand curves for land and seeds down. A new GM bean equilibrium is achieved with lower output than in case 1. Note that if the levy is large enough, then the inverse supply curve for other inputs shifts up far enough to negate the initial downward shift initiated by GM adoption. This results in production costs that are higher than in the NGM market, and GM bean production falls to zero.[19]

DC's welfare can be derived from the top panel of figure 1.6. If $\sigma < \chi(1-\varepsilon)$, then GM adoption is complete and welfare is given by area $acbd$. The amount of the levy paid to the GM innovator, area $efgh$, must be subtracted, however, from DC domestic welfare. If $\sigma > \chi(1-\varepsilon)$, then all bean production is NGM, and welfare is unchanged from the pre-adoption equilibrium. Note that DC welfare is always larger if $\sigma < \chi(1-\varepsilon)$; consumer welfare remains the same and bean output and producer surplus is larger.

The success of this policy is dependent on successful collection of the levy at one stage of the supply chain. Enforcing collection at the elevators or at the point of seed purchase is less costly for the IP innovator than is maintaining contracts with growers and monitoring enforcement of IP rights. This point is taken up in more detail in chapter 1.4.

[19] The same result can be shown in the case of a levy on seeds.

It would seem that the governments of developing countries have no incentive to negotiate such deals with foreign IP innovating firms. Domestic welfare benefits appear to be larger when GM seeds are pirated, with no proceeds paid to the innovating firm. However, as developed countries enter the WTO fold they are expected to fulfill their trade agreement obligations. The TRIPS agreement requires that all contracting parties enforce the IP rights of other contracting parties within the boundaries of their country. Failing to do so can lead to trade actions, as determined by the WTO's DSB. It is just such a trade action that IP innovating firms have as leverage when negotiating levies with developing country governments. By agreeing to implement a levy on black market IP, developing country governments hope to "buy off" IP innovators and avoid formal trade actions.[20]

1.3.5. Export Market Scenarios

The adoption of GM seeds may impact DC's export markets. The response of bean-importing nations to GM products may restrict GM imports or dictate the manner in which they are allowed to arrive. The following section sets up the framework in which the welfare consequences of various importer policies can be analysed. The black-market adoption pattern of case 1 is used as a baseline in the following scenarios. The effects of alternate adoption patterns combined with the following export market scenarios are explored in chapter 1.4.

<u>Case A - No Importer Response</u>

[20] The effects of retaliatory measures are considered in chapter 1.4.

This situation arises when consumers in bean importing nations have no preference for NGM beans over GM beans. The world price for GM beans remains the same as for NGM beans and there are no restrictions placed on DC's bean exports. The results of such a scenario are identical to those analysed in case 1. DC produces only GM beans and exports an amount of beans equal to ef in figure 1.3 at the prevailing world price of P_B^W. DC's welfare is equal to area $acbd$.

Case B - Pooled Equilibrium

Gaisford, et al. identify a situation in which there exist two types of consumers. Type A consumers prefer NGM to GM products and type B consumers are indifferent between products. Type A consumers acquire a larger marginal benefit from consuming NGM products than from GM products and have a correspondingly lower willingness-to-pay for GM products. If GM and NGM products are exported into the world market and sold together without separation or labelling, then the pooled product contains some GM and some NGM beans. As long as there exist some type A consumers, then the pooled product of combined GM and NGM beans is perceived as being of lower quality than pure NGM beans. The adverse effect on perceived quality reduces demand for the pooled product, and demand for beans shifts down. This result is similar to Akerlof's lemon analysis.

The "lemon" situation is represented by a decrease in world demand for beans. The downward shift in demand affects demand for both NGM and GM beans, since both are pooled together as a single product. Consumers do not know whether they are buying

NGM or GM beans, so the pooled product is viewed as being of lower quality, on aggregate. The size of the decrease in bean demand is a function of two factors. First, the larger is the share of type B consumers (who are indifferent between GM and NGM), the smaller is the demand shift. Second, the larger is the perceived proportion of the pooled product that is comprised of NGM beans, the smaller is the demand shift.

To derive the welfare effects of a pooled equilibrium, inverse world bean demand in equation (1.14) is modified as

$$P_B = \eta\delta\Omega - \alpha B, 0 < \eta < 1, 0 < \delta < 1. \tag{1.39}$$

The parameter η represents the portion of consumers that are made up by type Bs. The larger is the share of type Bs, the closer η is to one and the smaller is the demand shift. The parameter δ represents the share of the pooled product that is perceived to be NGM. The larger is the perceived share of NGM, the closer δ is to one and the smaller is the demand shift. The inequality constraints placed on η and δ are necessary for the existence of a pooled equilibrium. If both $\eta = 1$ and $\delta = 1$, then all consumers are type B and the final product is perceived to be made up entirely of NGM products - demand does not shift down. If these parameters are to equal one, then we revert to case 1.

Inverse world bean supply remains as in equation (1.15), and a new equilibrium level of bean consumption prevails at

$$B^{W*} = \frac{(\eta\delta\Omega - \theta)}{(\alpha + \beta)} \tag{1.40}$$

which is necessarily less than pre-GM bean output because the interaction effect of $\eta\delta$ is always less than one if there exist any type A consumers.

The new world bean price is

$$P_B^W = \eta\delta\Omega - \frac{\alpha(\eta\delta\Omega - \theta)}{(\alpha + \beta)}, \tag{1.41}$$

which is below the pre-GM world bean price because $\eta\delta < 1$.

DC's inverse GM bean supply function remains as in equation (1.23), which at the world price of equation (1.41) generates DC GM bean output of

$$\begin{aligned} B_{GM}^{DC*} &= \frac{\eta\delta\Omega - \frac{\alpha(\eta\delta\Omega - \theta)}{(\alpha + \beta)} - \varepsilon\chi - \gamma - \mu}{\lambda}. \\ &= \frac{P_B^W - \varepsilon\chi - \gamma - \mu}{\lambda} \end{aligned} \tag{1.42}$$

GM bean output is larger than the level of NGM output that would result in DC at the world price of equation (1.41), and GM adoption is complete. This result is intuitive since GM bean production is less costly than NGM bean production and both products can be sold at the same world price; no rational producer selects NGM seeds.

Figure 1.7 illustrates the effects of a pooled equilibrium on DC's bean industry. Only the GM market is included, since GM adoption is complete and NGM bean production falls to zero. World inverse bean demand shifts down from D_B^W to $D_B^W{}'$, resulting in a lower world bean price. DC produces GM beans in an amount equal to e at the new world price. DC welfare is equal to area $abcd$.

<u>Case C - Segregated Bean Markets</u>

Case C considers situations in which GM and NGM beans are produced along two distinct supply chains[21] and sold separately in the world market, each according to its own demand function. The world bean market consists of type A and type B consumers, with type A consumers willing to pay more for NGM beans. Type A consumers prefer NGM products, so NGM beans are perceived as higher-quality products than are GM beans. The price of NGM beans therefore always includes a quality premium over GM beans. As long as the price of NGM beans is higher than the price of GM beans, type B consumers buy only GM beans. The present analysis assumes that DC consumers are all type B, though this assumption could be changed within the context of the model.

i) Segregation without trade action

[21] It is shown below that DC produces only GM or NGM beans, and not both. Other countries, however, may produce either or both so that the world bean market contains both GM and NGM products.

We first consider the scenario wherein DC's export markets do not limit imports of either GM or NGM beans. Importing-country governments do not interfere in the bean market beyond imposing labelling requirements on GM and NGM producers.[22] The acceptance of GM products in the market is dependent on consumer preferences. A methodology similar to Gaisford, et al. is applied to the fixed proportions bean model in the forthcoming analysis[23].

The introduction of GM beans has two initial effects. First, since there exist type B consumers who are willing to substitute GM beans, the demand for NGM beans shifts down. World inverse demand for NGM beans is represented as

$$\psi\Omega - \alpha B, 0 < \psi < 1 . \tag{1.43}$$

Equation (1.43) differs from world inverse bean demand in equation (1.14) by the parameter ψ, which shifts down the intercept of the inverse demand function. The larger is the proportion of type A consumers, the larger is ψ. Assuming that DC is the only producer of the GM product, the world supply of NGM beans remains as in equation (1.15). A new NGM world equilibrium generates NGM bean output and price of

[22] The ability of developing countries to implement segregation and labelling programs, are discussed in chapter 1.5.
[23] Gaisford, et al. analyse the comparative statics of the segmentation of a market for a single product into two distinct markets; one for GM and another for NGM.

$$B_{NGM}^{W*} = \frac{\psi\Omega - \theta}{(\alpha - \beta)} \tag{1.44}$$

and

$$P_{B,NGM}^{W*} = \psi\Omega - \frac{\alpha(\psi\Omega - \theta)}{(\alpha + \beta)}. \tag{1.45}$$

Note that NGM bean output and price are lower than pre-GM bean output and price because $\psi < 1$.

The existence of type B consumers gives rise to an inverse demand function for GM beans, which is given by

$$P_{B,GM} = \tau\Omega - \alpha B, 0 < \tau < \psi. \tag{1.46}$$

DC's supply of GM beans is given by equation (1.23). Two important points must be made about the GM market. First, the restriction that $\tau < \psi$. This inequality states that the inverse demand function for GM beans must be below the inverse demand function for NGM beans; the GM product is a *weakly inferior innovation* in the preference ordering of consumers (Lapan and Moschini). No consumers prefer GM to NGM beans, but type A consumers prefer NGM to GM beans. Therefore, consumers never pay more for GM beans than for NGM beans.

Second, the inverse demand function for GM beans is perfectly elastic over some range of output because demand for GM beans is conditional on the equilibrium price for NGM beans (Gaisford, et al.). Type B consumers are never willing to pay more for GM beans than for NGM beans, so demand for GM beans above the equilibrium price for NGM beans is zero. Therefore, world demand for GM beans is kinked. The level of output at which the kink occurs can be determined by evaluating the level of GM output where the NGM bean price intersects the GM bean demand curve. This occurs at

$$\begin{aligned} B^{\Diamond} &= \frac{\Omega(\tau-\psi)+\dfrac{\alpha(\psi\Omega-\theta)}{(\alpha+\beta)}}{\alpha} \\ &= \frac{\tau\Omega-P_{B,NGM}^{W*}}{(\alpha+\beta)} \end{aligned}. \tag{1.47}$$

To the left of $B^{\Diamond}$, world inverse GM bean demand is perfectly elastic at $P_{B,NGM}^{W*}$. To the right of $B^{\Diamond}$, world inverse GM bean demand is as shown in equation (1.46).

The GM bean market equilibrium is given by the intersection of GM bean supply (equation (1.23)) and GM bean demand (equation (1.46)), and generates price and output of

$$P_{B,GM}^{W*} = \tau\Omega - \frac{\alpha(\tau\Omega-\varepsilon\chi-\gamma-\mu)}{(\alpha+\lambda)} \tag{1.48}$$

and

$$B_{GM}^{W*} = \frac{P_{B,GM}^{W*} - \varepsilon\chi - \gamma - \mu}{\lambda}. \tag{1.49}$$

The difference between the world NGM and GM bean prices dictates whether DC farmers adopt GM technology in the case of segmented markets; the price differential must be less than the GM cost savings for DC farmers to have any incentive to adopt. That is, $P_{B,NGM}^{W*}$ from equation (1.45) minus $P_{B,GM}^{W*}$ from equation (1.48) must be less than $\chi(1-\varepsilon)$. If the price differential is sufficiently small, then the cost savings of GM seed technology are sufficient to warrant adoption. If not, then no farmer has an economic incentive to adopt GM technology.

If the price differential is sufficiently small and DC farmers adopt GM technology, then DC produces beans (all of them GM) in an amount equal to

$$B_{GM}^{DC*} = \frac{P_{B,GM}^{W*} - \varepsilon\chi - \gamma - \mu}{\lambda}. \tag{1.50}$$

If the NGM-GM bean price differential is sufficiently large, then DC produces only NGM beans, in an amount equal to

$$B_{NGM}^{DC*} = \frac{P_{B,NGM}^{W*} - \chi - \gamma - \mu}{\lambda}. \tag{1.51}$$

The price differential between GM and NGM beans depends on the ratio of type A to type B consumers. If most consumers are type A, then demand for the substitute GM product is low. The shift parameter in the NGM demand function, ψ, is close to one and NGM inverse demand may not shift far enough to induce DC farmers to adopt GM seeds. The post-innovation price for NGM beans remains near its pre-GM level. Correspondingly, the demand for GM beans is low (τ in equation (1.46) is small) and the price of GM beans is relatively low. The price differential is large and DC produces only NGM beans. If, however, the proportion of type B consumers is large, then the price differential is small and DC adopts GM bean technology. Recall that the price differential can never be negative - as long as there exist some type A consumers, the price of NGM beans is always above the price of GM beans.

The results of case C.i are dependent on there being type A consumers who do not switch to GM products regardless of the price differential between NGM and GM products. As discussed above, a larger share of consumers that is willing to switch to GM products generates a smaller price differential between NGM and GM products. Contrarily, a smaller share that is willing to shift generates a larger price differential. If type A consumers are willing to switch to GM products when the price differential is large enough, then an unstable equilibrium results. However, as long as type A consumers do not switch to GM beans regardless of the price differential, a stable equilibrium can be attained. Gaisford, et al. address the situation of inframarginal consumers in more detail.

ii) Segregation with an importer ban on GM beans

The case in which some or all of DC's export markets impose a ban on GM beans can be analysed using the framework of the previous scenario. A government that bans GM imports overrides consumer preferences as represented by parameters ψ and τ; the nation becomes *de facto* exclusively type A. If DC exports beans to only one country, and that country bans GM beans, then the world demand curve facing DC remains as in equation (1.14) and the world demand for DC's GM beans falls to zero. If DC adopted GM seeds prior to the import ban, then it loses its entire GM export market and has only domestic type B consumers to supply[24]

The situation in which only some of DC's export markets ban GM products can be understood through various relative values of ψ and τ. By banning GM imports, those nations that restrict GM imports become exclusively type A, thus bringing the global value of ψ closer to one and the value of τ nearer zero.

1.4. Discussion

The model in chapter 1.3 provides tools that can be used to analyse a range of policy scenarios. This chapter combines adoption scenarios (cases one through three) with export market scenarios (cases A through C) and discusses the welfare effects and policy options. Table 1.1 illustrates possible adoption policy and export market scenario combinations. The outcomes in table 1.1 illustrate probable outcomes, but recall that results are dependent on model parameters (i.e. complete adoption in the case

[24] The possibility that adoption of GM technology is irreversible is discussed in chapter 1.4.

of an IP levy is dependent on the size of the levy). The remainder of this chapter discusses selected scenario combinations in detail. Two combinations are selected for closer analysis because of their relevance to current developing country situations. The adoption and export market combinations of (case 1 + case A) and (case 3 + case C.ii) are discussed below.

Table 1.1 - Policy Combination Scenarios

	Pirate GM seeds	Monopoly, drastic	Monopoly, nondrastic	IP Levy
No Export Market Response	Scenario 1 (cases 1 + A)	Complete adoption (cases 2 + A)	Complete adoption (cases 2 + A)	Adoption dependent on levy (cases 3 + A)
Pooled Market	Complete adoption (cases 1 + B)	Complete adoption (cases 2 + B)	Complete adoption (cases 2 + B)	Adoption dependent on levy (cases 3 + B)
Segregated Market - No Import Ban	Adoption dependent on NGM-GM price differential (cases 1 + C.i)	Adoption dependent on NGM-GM price differential (cases 2 + C.i)	Adoption dependent on NGM-GM price differential (cases 2 + C.i)	Adoption dependent on levy (cases 3 + C.i)
Segregated Market - Import Ban	Adoption dependent on NGM-GM price differential (cases 1 + C.ii)	Adoption dependent on NGM-GM price differential (cases 2 + C.ii)	Adoption dependent on NGM-GM price differential (cases 2 + C.ii)	Scenario 2 (cases 3 + C.ii)

A few important points are worthy of mention before proceeding. First, there is no general ordinal ranking of policy alternatives. The economic surplus methods of chapter 1.3 generate welfare results that are dependent on demand and supply parameters. Welfare measures could be compared between scenarios if empirical estimates of the relevant demand and supply parameters were available for a case study. The forthcoming discussion of various scenarios attempts to provide guidance on the possible size of welfare effects in developing countries.

The model in chapter 1.3 is presented in an order that follows the soybean supply chain in chronological order. The GM adoption decision and production patterns, and their resultant welfare effects, are presented first. Welfare effects of various export market scenarios are presented second. DC's policy analysis should, however, begin in stage two. DC should form an expectation of likely export market effects before deciding on an adoption policy that maximises domestic welfare. For example, if DC is likely to face a devastating loss of export markets in the event of GM adoption, then the best method of GM adoption may be moot - DC should consider not adopting GM crops. Contrarily, if DC anticipates that its export markets will not express any adverse reaction to GM products, then the most appropriate policy may be to fully adopt GM technology so as to best take advantage of the production cost advantage.

Scenario 1 - Pirate seeds, IP Innovator Retaliation and no Importer Response

Scenario one analyses the combination of case one and case A, with the added component of retaliatory trade measures by the host country of the IP innovator. Recall that the TRIPS agreement allows for cross-retaliation under WTO agreements if a member country is found in violation of its TRIPS agreement obligations. The efficacy of cross-retaliation as a means of inducing compliance with the TRIPS agreement has been called into question (Yampoin and Kerr); the incentives to protect foreign IP rights have been shown to decline with the size of the pirate industry and the costs of enforcing IP rights. Despite this uncertainty, it is the threat of retaliatory trade measures that forms the basis of case three, the levy on pirated GM technology. Developing countries' only incentive to agree to a levy is the avoidance of TRIPS cross retaliation. For this reason, the effects of retaliatory trade measures warrant attention.

The starting point for analysing this scenario is case one, where GM technology is pirated and adoption is complete. Welfare gains are initially as illustrated in figure 1.3; consumer surplus of acP_W and producer surplus of $P_W bd$. If the host country of the IP innovator pursues its case with the WTO and wins a ruling from the DSB, then retaliatory trade measures are allowed in an amount equal to the innovator's trade lost due to DC's pirate industry (WTO). Calculating the amount of lost trade, however, presents a practical problem. To measure the amount of loss suffered by the IP innovator, a counter-factual calculation is required (Yampoin and Kerr). If DC pirated GM technology, then the IP innovator was never provided the opportunity to sell its product at the monopoly price. Estimated demand and marginal revenue functions are required to ascertain what the seed market equilibrium would have been in the case of a monopoly. We turn to case two and figures 1.4 and 1.5 for this information. The bottom panels in figures 1.4 and 1.5 illustrates the case in which the IP innovator monopolises the GM seed market for drastic and nondrastic innovations, respectively. The counter-factual lost trade to the IP innovator is area $P_S^M fg0$ in each case. Note that this is not equal to area $P_S^{*}{}'ij0$ from figure 1.3.

The IP innovator's host country can therefore initiate trade actions that penalise DC in an amount equal to $P_S^M fg0$. The cross-retaliation nature of the TRIPS agreement suggests that this action will come in the form of tariffs on products imported from DC. The complainant country is authorised to choose the products upon which tariffs are placed. While it is possible that such tariffs would be applied to DC beans, products selected for retaliation are typically chosen for their political sensitivity and could well be applied to other imports from DC. Therefore, retaliatory trade actions may not affect DC's soybean industry. In making a policy decision on adoption, however, DC must be aware that the loss of area $P_S^M fg0$ must be subtracted from welfare gains accrued in the bean industry.

Scenario 2 - IP Levy and an Import Ban

The scenario warrants attention for two reasons. First, if DC completely adopts GM crops (as the model in chapter 1.3 predicts for some circumstances) and its consumer nation(s) bans GM imports, then all of DC's bean exports are banned. Second, a levy on GM technology is a topical issue for several South American countries. Brazil and Paraguay's recent agreement to enforce a levy on soybean seeds (The Western Producer) indicates that the adopting countries believe there to be benefits of such a deal. Brazil and Paraguay likely recognise the potential threat of trade action by the US in retaliation for not enforcing Monsanto's IP rights. At the same time, policy-makers are faced with the difficult reality of enforcing IP rights in their own countries. Gaisford, et al. point out that enforcement of IP rights involves two types of costs.

First, there are the costs of identifying and monitoring pirate firms. Second are the costs of ensuring the efficacy of the IP rights protection system. These costs include the maintenance of a legal system and of offsetting corruption in the monitoring system. The second type of cost is likely to be particularly high in developing countries. The institutions that facilitate an effective legal system are underdeveloped in developing countries (Hobbs and Kerr) and rampant corruption increases the costs of bureaucratic oversight (Alam). Brazil and Paraguay may believe that the potential costs of enforcing Monsanto's IP rights exceed the benefits of GM adoption. Negotiating a check-off style levy with Monsanto may prove to be a policy that avoids the transaction costs of enforcing IP rights while averting retaliatory trade action from the US.

A levy is also an appealing policy from a developing country's perspective because a portion of the proceeds is likely to be channelled into domestic public research projects (Traxler). The negotiating government holds bargaining power because a levy is negotiated with, rather than imposed by, the IP innovator. Traxler believes that some of this bargaining power will be used to retain a portion of levy proceeds in the developing country to increase research capacity. The goal of such research would be to either decrease reliance on foreign biotechnology research or to develop breeding programs that produce GM crop varieties better suited to local agronomic conditions, as discussed in chapter 1.2.1.

To analyse scenario two, we begin at the equilibrium result of case three, where DC fully adopts GM crops and produces beans according to equation (1.37). If DC loses all of its export market to a ban, then there are two possibilities. If the adoption of GM crop technology is reversible, then DC abandons all GM production and reverts back to its pre-GM equilibrium. Foreign demand for DC's GM beans falls to zero, pushing the world price of GM beans to zero. There is no incentive to produce GM beans so all bean production in DC reverts back to NGM.[25] A comparison of static welfare measures from DC's pre-GM situation to the post-adoption, post-ban situation reveals no change in domestic welfare; both are equal to area *acbd* in figure 1.2. Such a comparison obviously overlooks sizable adjustment costs that would be incurred by adopting and then "un-adopting" GM crop technology. The costs of a wholesale change in crop planting would likely be very large.

Another, and perhaps more realistic, possibility in scenario three is that the "un-adoption" of GM crops is prohibitively expensive, making the adoption decision irreversible. Once GM seeds are released, a regulation that bans their use could be difficult, or even impossible, to enforce (Gray, McNaugton and Stovin). The same factors that make protection of IP rights expensive in developing countries would make enforcement of a ban on an already-released GM product extremely difficult. Also, developing country agriculture is characterised by a large number of farms who produce for consumption at home. Even if the total loss of GM export markets reduced the price of GM beans to zero, subsistence farms would have no incentive to give up the more cost-efficient GM product.

[25] The exception to this case is if the cost advantage provided by the GM trait is so large that it is still more profitable to produce GM beans for domestic consumption and then allocate remaining resources to NGM production for foreign markets.

If GM adoption is irreversible and DC faces a foreign ban on its products, then a sizable welfare loss occurs. The top panel of figure 1.6 (pirate seeds with a levy) is reproduced in figure 1.8 with only the required curves for the current analysis. After adopting GM seeds, and before a ban, DC's welfare is equal to consumer surplus of area $abP_{B,GM}^{W*}$ plus producer surplus of area $P_{B,GM}^{W*}cd$. A foreign ban on DC's bean exports reduces world GM bean demand to coincide with DC's domestic demand for GM beans, and the equilibrium price for GM beans falls to $P_{B,GM}^{W}{}'$. DC's consumer welfare increases (because of the lower price for GM beans) to area $aeP_{B,GM}^{W}{}'$ but producer surplus falls to area $P_{B,GM}^{W}{}'ed$. This scenario highlights the importance of DC forming an expectation of probable export market responses to its adoption of GM technology. If a ban is likely, then DC clearly benefits from not adopting GM technology at all.

The political economy of trade agreements plays an important role in the welfare analyses of these scenarios. Scenario two involves trade actions that are governed by the TBT and SPS agreements. If an import ban cannot be justified under these agreements, then importers are in violation of their WTO obligations. Though the WTO is intended to provide protection from such violations, adopting countries should be prepared to face such violations. Member countries appear ready and willing to violate their WTO obligations in response to political pressure (Kerr and Hobbs).

1.5. Conclusions

The decision to adopt genetically-modified crop technology is complicated by potential trade effects, particularly in developing countries. The benefits of lower production costs, higher nutritional value and hardier crops must be weighed against the possible negative welfare effects of trade actions by other countries.

The majority of agricultural biotechnology is created in developed countries, so that the intellectual property rights embodied in genetically-modified crops are not held in developing countries. Developing countries must be aware of the potential for retaliatory trade measures if foreign-owned intellectual property is pirated. A promising method for dealing with the complications of enforcing intellectual property rights is the imposition of a levy on production of genetically-modified crops. A levy could be particularly beneficial in developing countries for three reasons. a levy could work to pacify innovating firms whose intellectual property is pirated, thereby avoiding retaliatory trade measures. Second, a check-off style levy is likely to involve lower transaction costs than the type of contracts that seed companies generally utilise in developed countries. Finally, some of the proceeds from a levy could be diverted into local agricultural research to promote the type of varietal improvements that are discussed by Evenson.

Consumers and civil society groups in several nations have reacted adversely to genetically-modified products and have pressured their respective governments to impose restrictions on the use and trade of genetically-modified foods. Developing countries should be aware of the effects of such restrictions (which could include import bans and identity-preservation system requirements) and factor these effects into their policy decisions. A lost export market could be devastating enough to more than negate any positive gains from cost-reducing genetically-modified technology. Developing countries must be particularly aware of such potential losses if the decision to adopt genetically-modified crop technology is irreversible. It is important to conduct a welfare analysis before adoption occurs.

This essay demonstrates that a developing country will either entirely adopt GM technology, or will not adopt at all.[26] As such, an identity preservation system, and a developing country's (in)ability to introduce a system, is not important in maintaining export markets.

The policy decision to adopt genetically-modified crops may be simpler for developing countries that produce food primarily for domestic consumption. Such countries have less to lose in international markets, so the efficiency and nutritional benefits of genetically-modified technology might dominate potential losses. However, those countries that rely heavily on international trade in agricultural products should be aware of the potential downside risks of adopting genetically-modified crop technology.

[26] This result is dependent on all farmers being equally suited to the GM trait. The prospect of heterogeneous farmers is addressed below.

A natural extension of this essay is a case study that uses empirical estimates of the relevant demand and supply functions. Welfare effects could be calculated and used in a comparison of policy alternatives.

Another interesting extension would be to model adopting-country farmers as being heterogeneously suited to the GM technology.[27] Those farmers best-suited to GM technology would adopt first, and an equilibrium may emerge in which GM adoption is not complete. Such a result would have interesting welfare implications, and would introduce the possible complication of DC's (in)ability to institute an identity preservation system. The estimated costs of identity-preservation systems are high, and developing countries likely do not have the required institutional mechanisms to facilitate such systems. Even in scenarios which do not result in full adoption of genetically-modified crops, developing countries that cannot implement a successful identity-preservation system become *de facto* comprised entirely of genetically-modified crops. On a similar note, even if policy makers decide to not allow adoption of genetically-modified crops in a developing country, it is uncertain if the institutional and regulatory capacity exists to enforce such a decision. Proving that one's country does not use GM technology may be an onerous, if not impossible, task for some developing countries.

[27] The issue of producer heterogeneity has been addressed in Fulton and Giannakas and in Malla and Gray.

Figure 1.1. Fixed Proportions Bean Industry

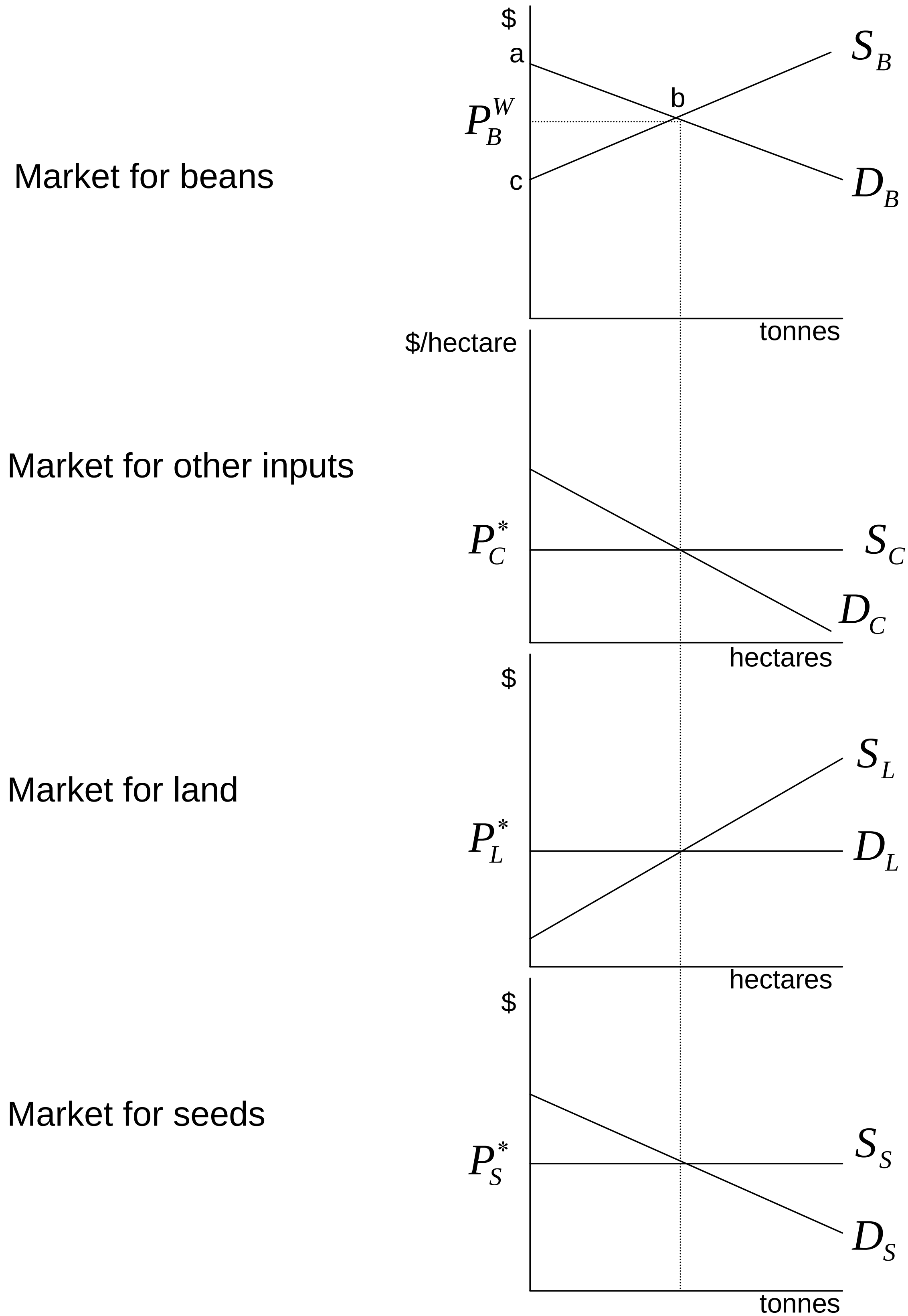


Figure 1.2. Fixed Proportions with International Trade

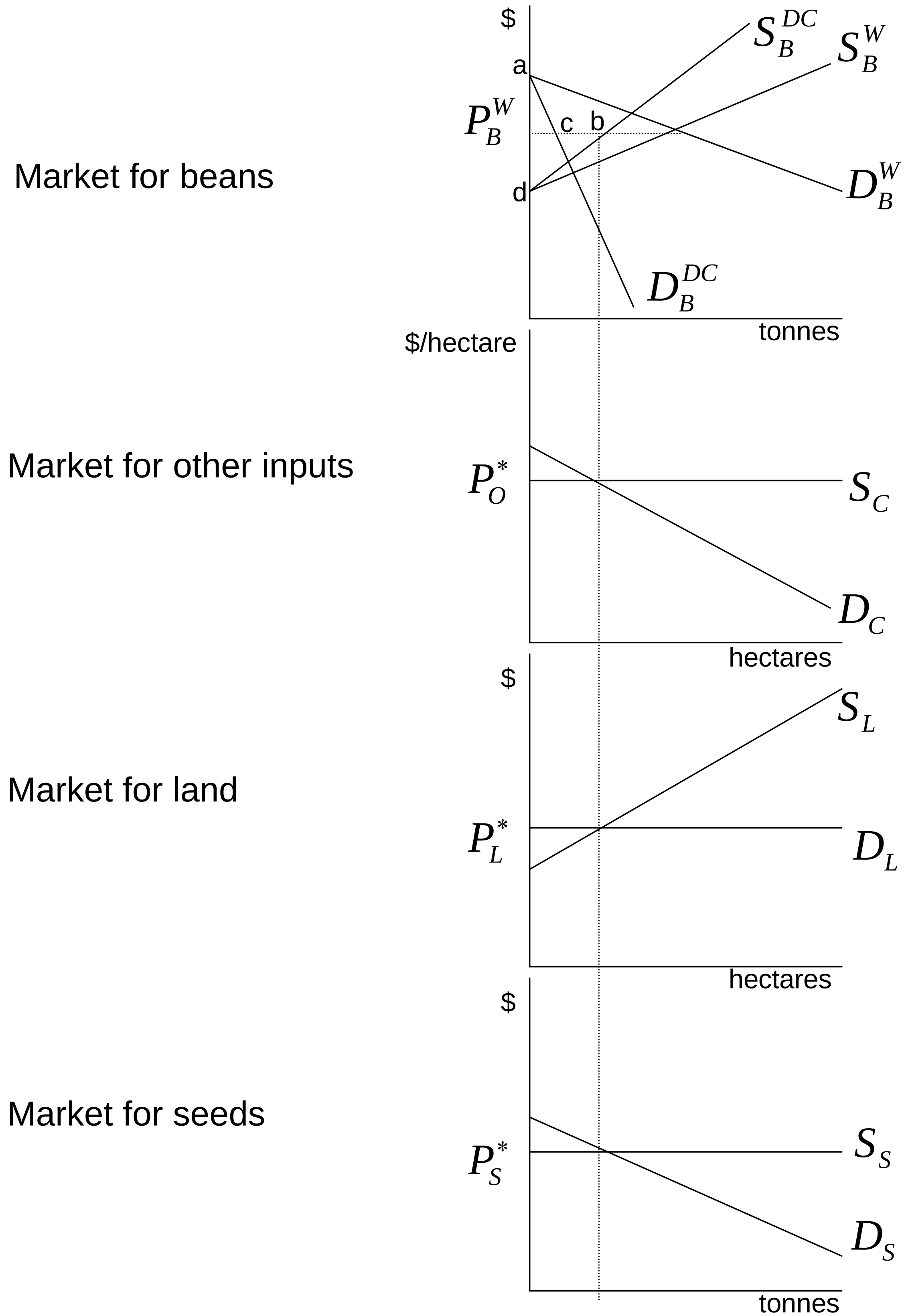


Figure 1.3. Fixed Proportions with Technological Change

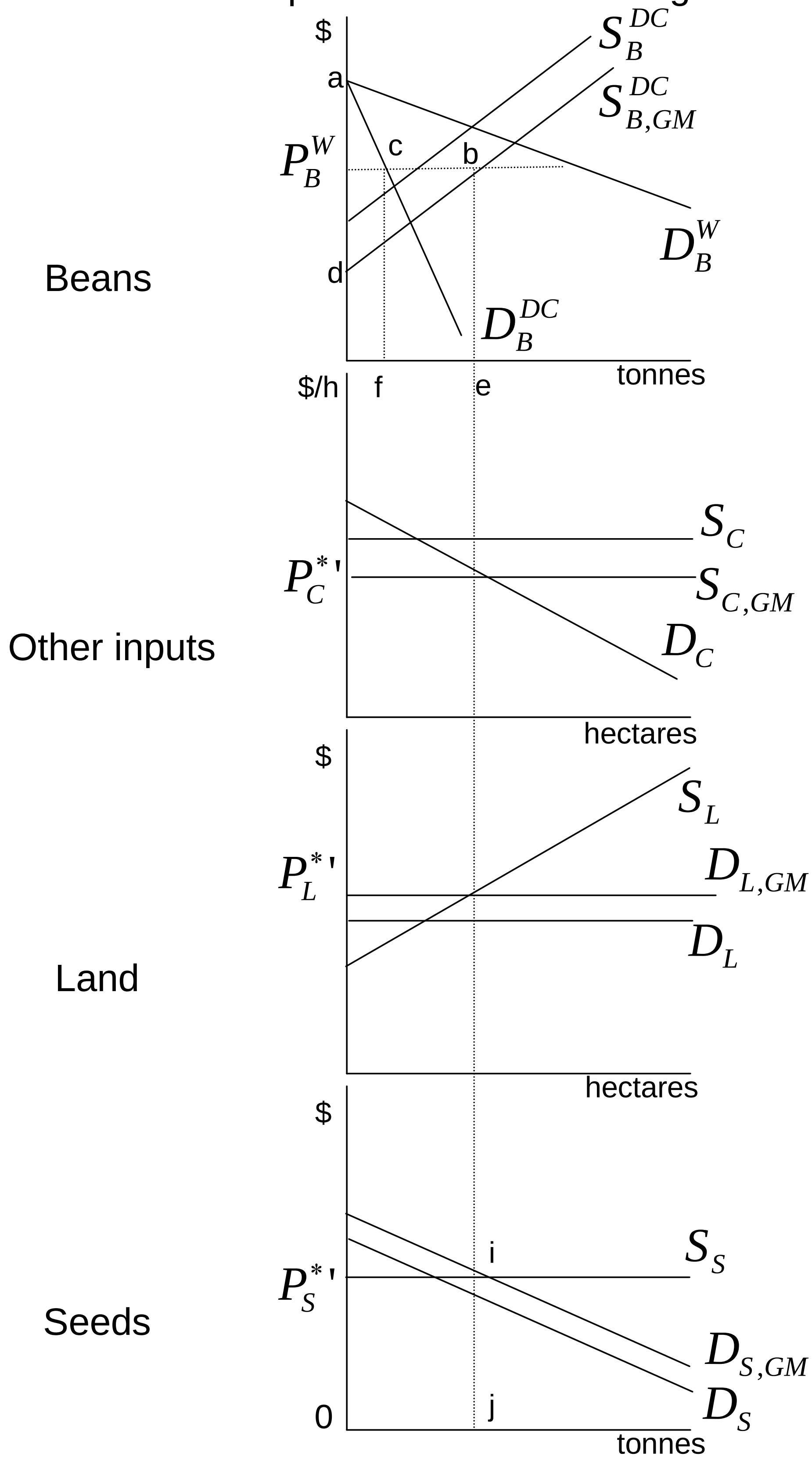


Figure 1.4. Monopoly - drastic innovation

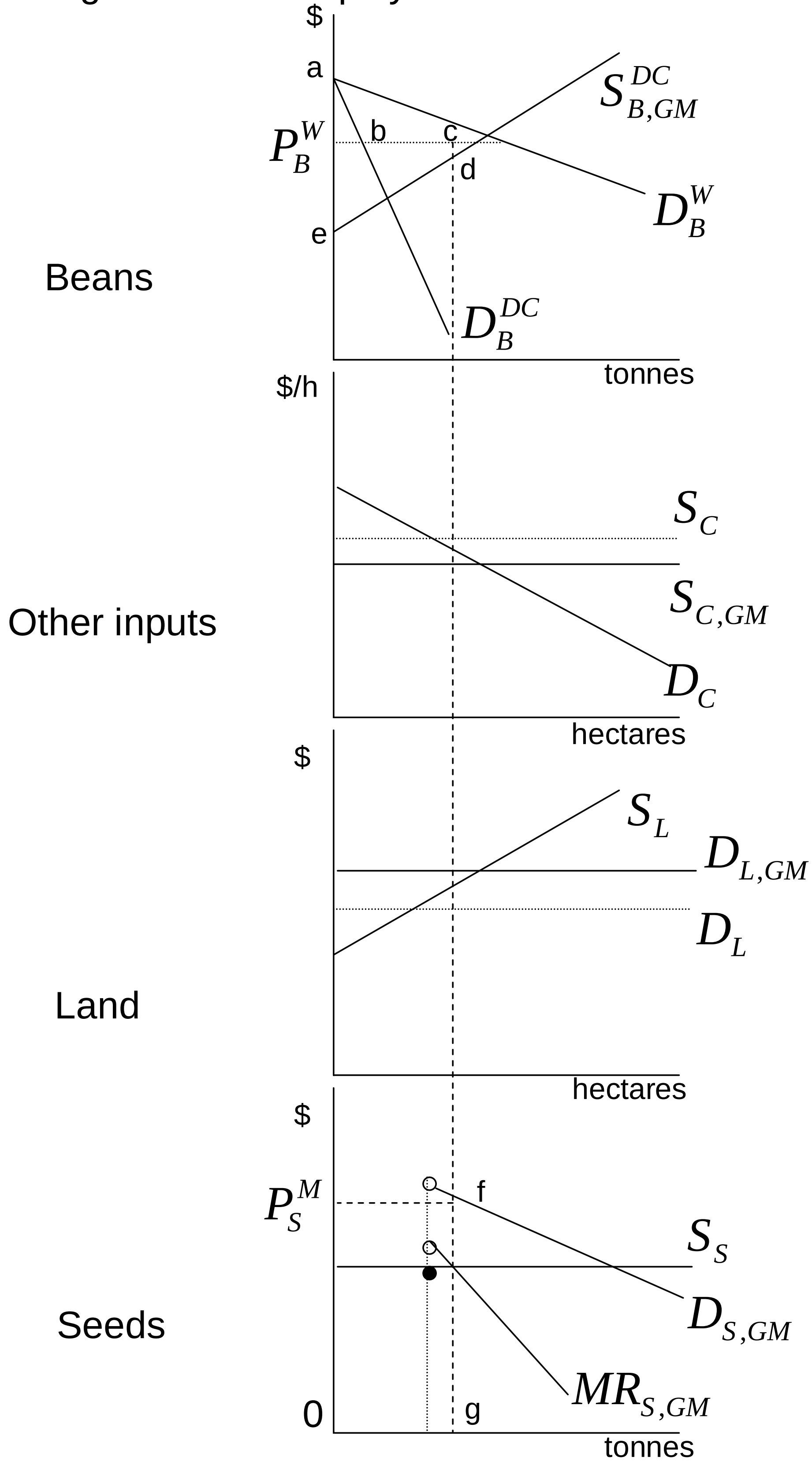


Figure 1.5. Monopoly - non-drastic innovation

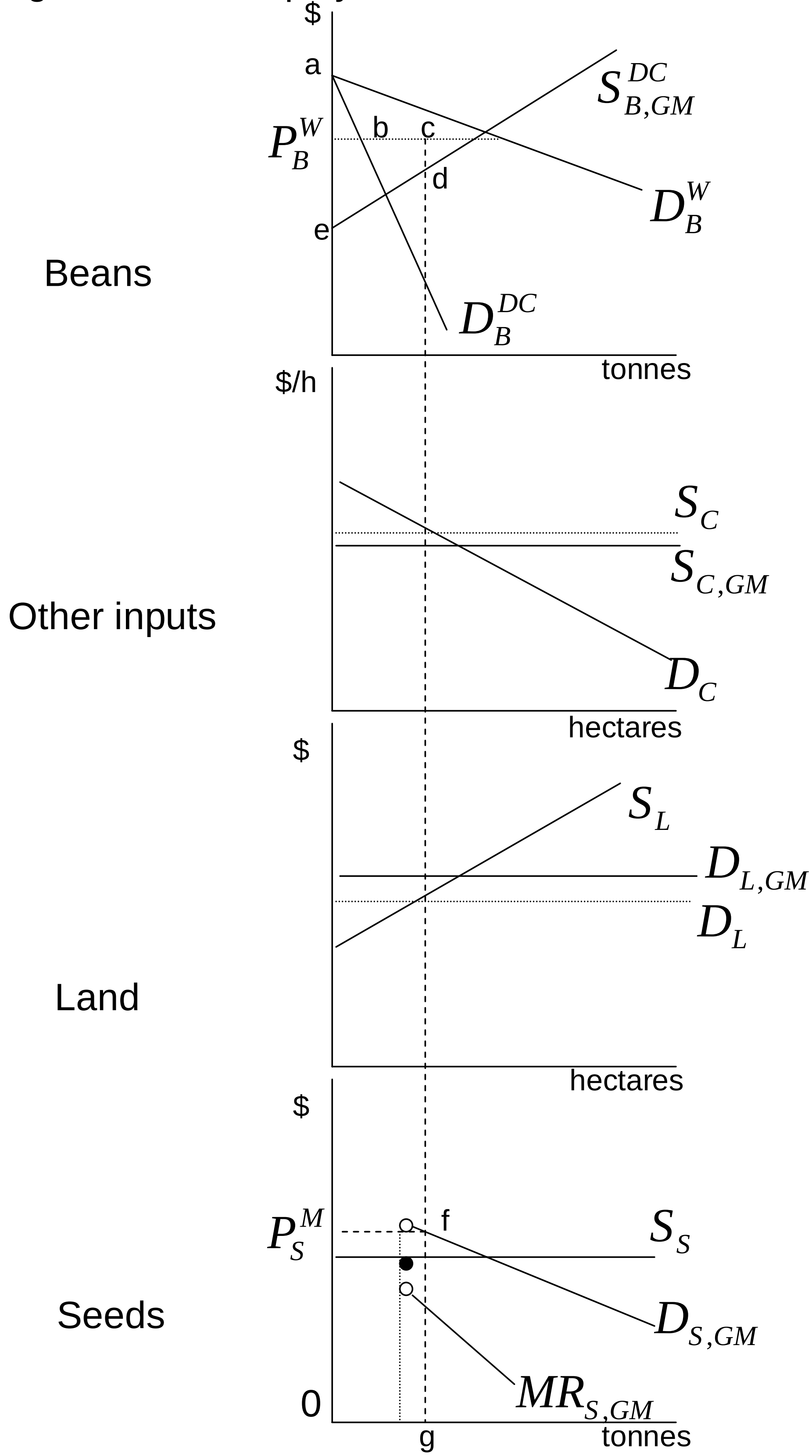


Figure 1.6. Fixed Proportions with Elevator Levy

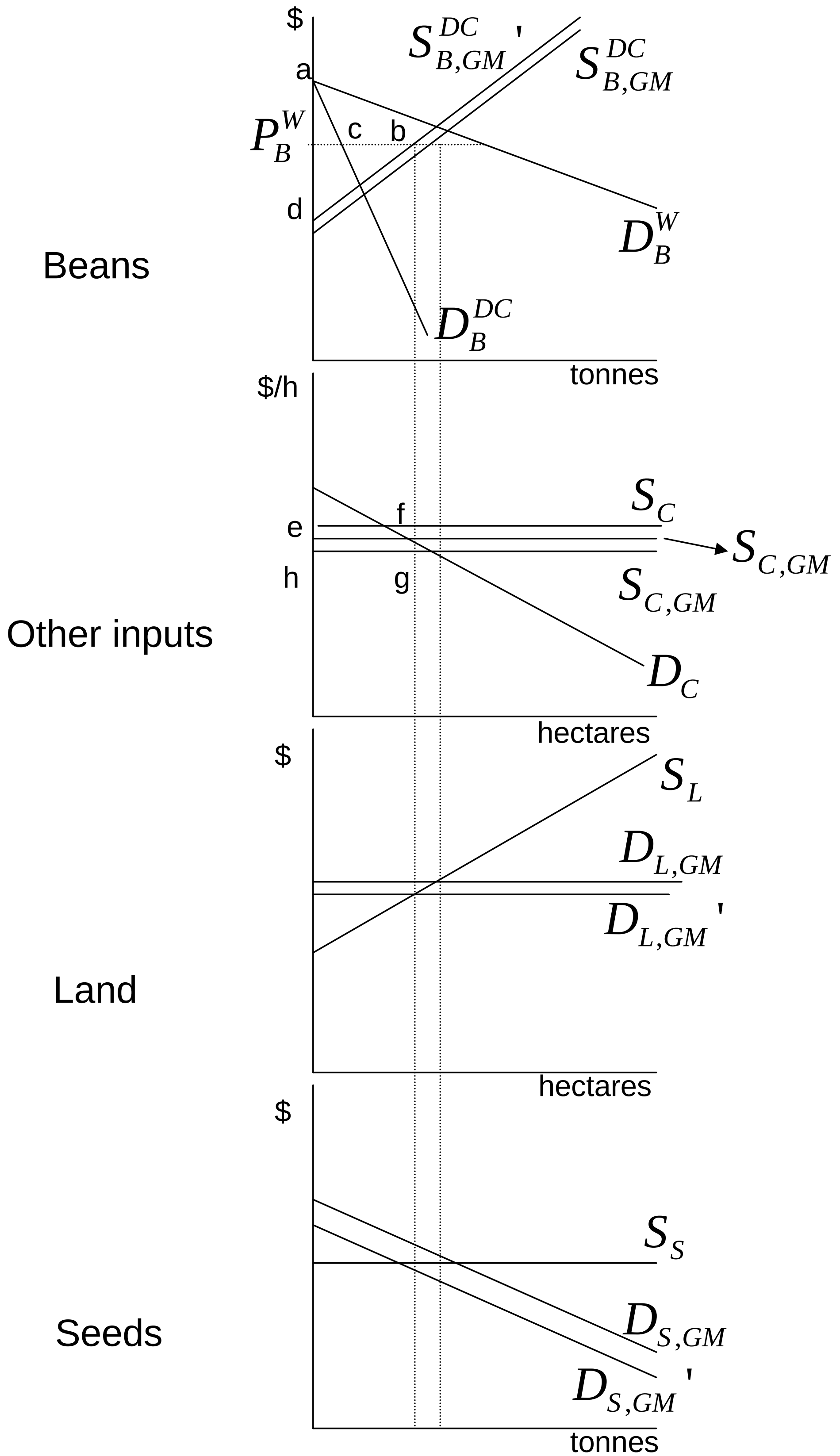


Figure 1.7. Fixed Proportions with Pooled Equilibrium

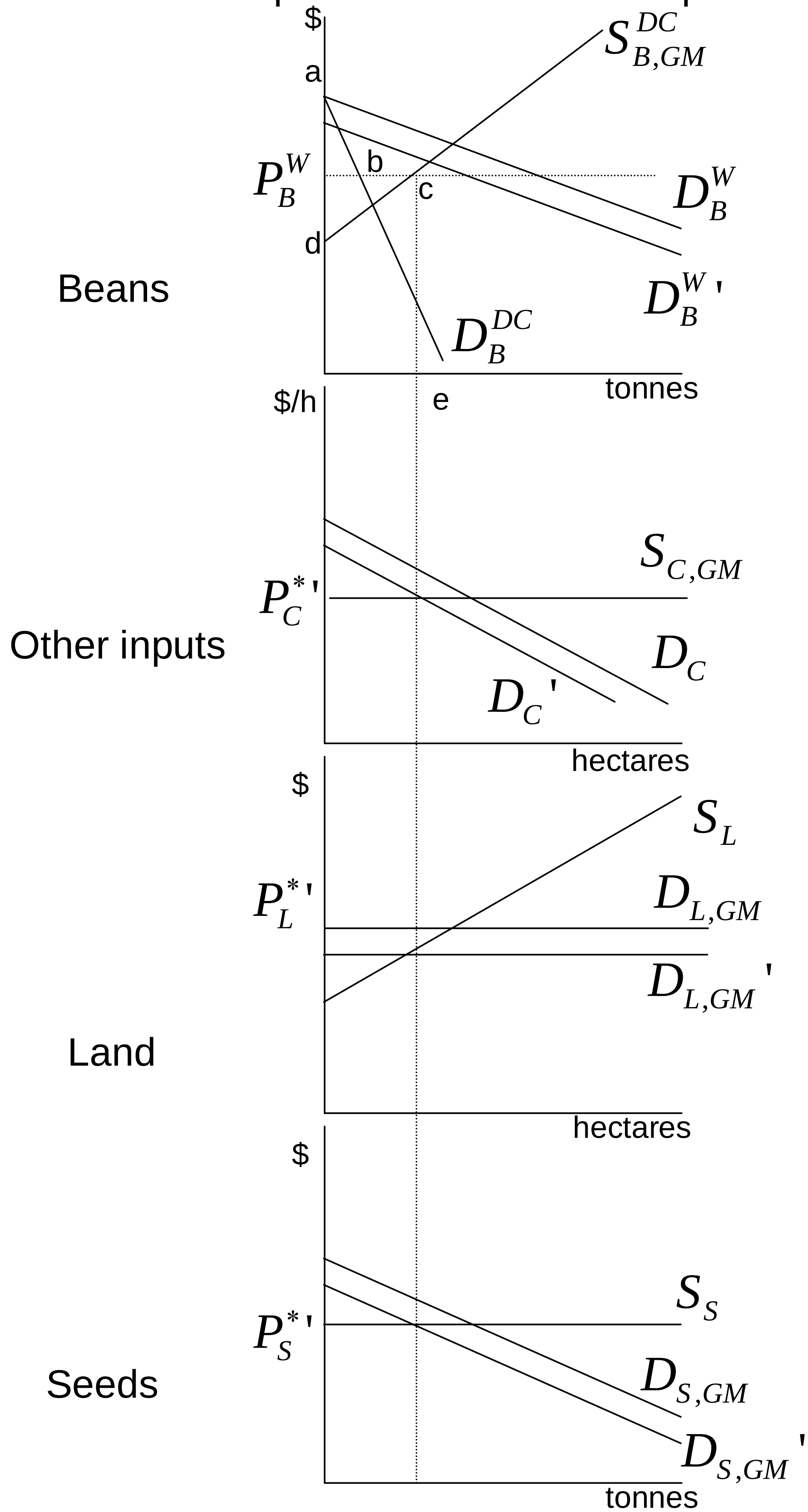
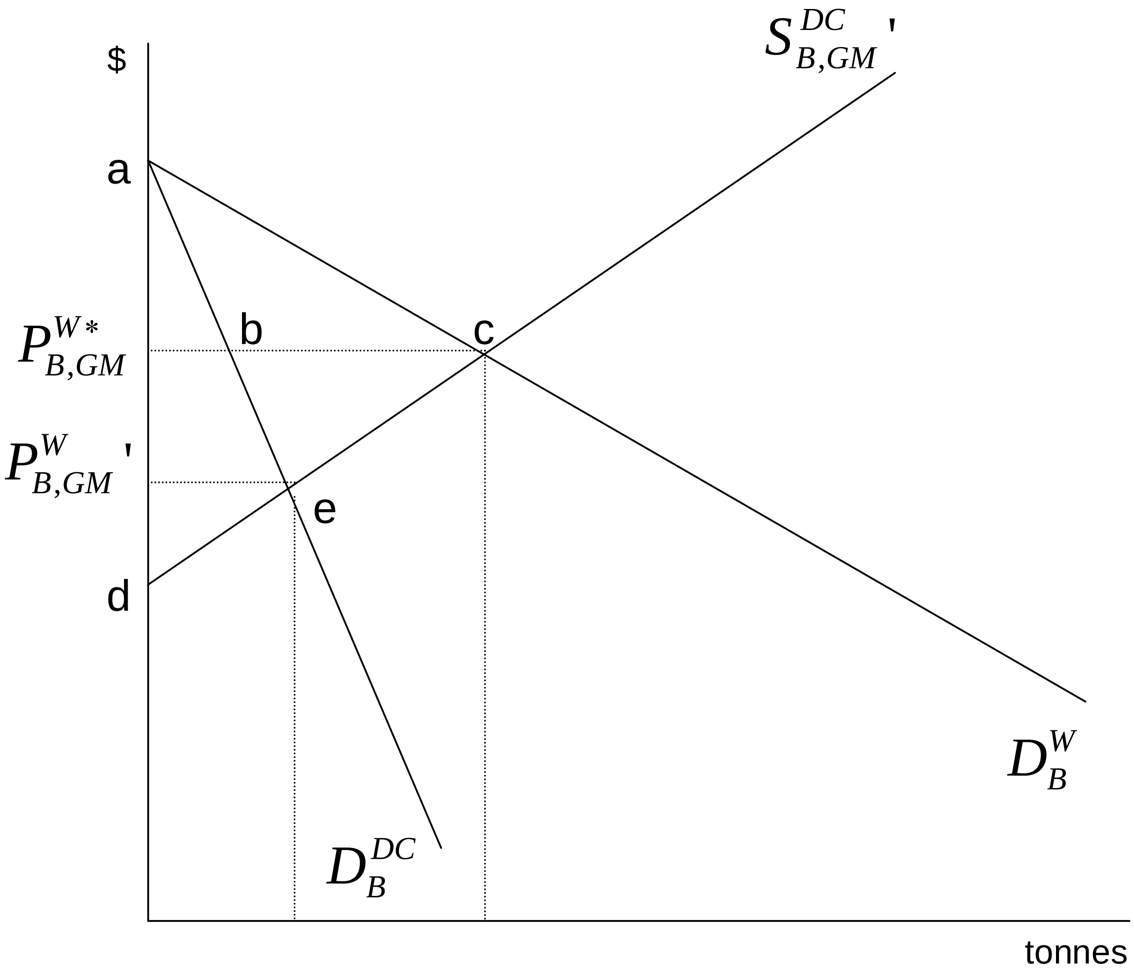


Figure 1.8. Scenario 2



REFERENCES

Akerlof, G.A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84, no. 3(1970): 488-500.

Alam, M.S. "Some Economic Costs of Corruption in LDCs." *Journal of Development Studies* 27, no. 1(1990): 89-97.

Alston, J., G. Norton and P. Pardey. *Science under scarcity: principles and practice for agricultural research evaluation and priority setting*. Ithaca: Cornell University Press, 1995.

Bullock, D.S. and M. Desquilbet. "The economics of non-GMO segregation and identity preservation." *Food Policy* 27(2002): 81-99.

Dahlsten, S. "Agriculture - Key Issues for Policy Coherence for Development." OECD, 2004. http://www.oecd.org/dataoecd/22/53/25507214.pdf.

Evenson, R. "GMOs: Prospects for Productivity Increases in Developing Countries." *Journal of Agricultural & Food Industrial Organization* 2, no. 2 (2004).

Gaisford, J., J. Hobbs, W. Kerr, N. Perdikis and M. Plunkett. *The Economics of Biotechnology*, Northhampton: Edward Elgar Press, 2001.

Fulton, M. and K. Giannakas. "Inserting GM Products into the Food Chain: The Market and Welfare Effects of Different Labeling and Regulatory Regimes." *American Journal of Agricultural Economics* 86, no.1 (2004).

Gray, R., J. McNaughton, and D. Stovin. 2001. "Biotechnology and Lesser Developed Countries: An Overview of the Issues" in *Transforming Agriculture: The Benefits and Costs of Genetically Modified Crops*, University of Saskatchewan, p. 119-130.

Haggui, F. "Cost of EU Opposition to Genetically Modified Wheat in Terms of Global Food Security", PhD thesis, University of Saskatchewan, 2004.

Hobbs, J. and W. Kerr. "Transaction costs" *The Current State of Economic Science* 4 (1999): 2111-33, edited by S. Bhagwan Dahiya and V. Singh.

Hobbs, J., W. Kerr and P. Phillips. "Identity Preservation and International Trade: Signalling Quality Across National Boundaries." *Canadian Journal of Agricultural Economics* 49 (2001): 567-579.

Isaac, G. W. Kerr, W. "Genetically Modified Organisms at the World Trade Organisation: A Harvest of Trouble", *Journal of World Trade* 37, no.6 (2003): 1083-1095.

Isaac, G., M. Phillipson, and W. Kerr (2002) *International Regulation of Trade in the Products of Biotechnology*. Estey Centre Research Papers, No. 2. Saskatoon: Estey Centre for Law and Economics in International Trade.

James, C. 2003. “Global Status of Commercialized Transgenic Crops: 2003.” ISAAA Briefs No. 30. ISAAA: Ithaca.

James, C. 2004. “Preview: Global Status of Commercialized Biotech/GM Crops: 2004.” ISAAA Briefs No. 32. ISAAA: Ithaca.

Kalaitzandonakes, N., R. Malstberger, and J. Barnes. “Global Identity Preservation Costs in Agricultural Supply Chains.” *Canadian Journal of Agricultural Economics* 49(2001): 605-615.

Kerr, W. “Science-based Rules of Trade - A Mantra for Some, An Anathema for Others.” *Estey Centre Journal Of International Law and Trade Policy* 4, no. 2(2003): 86-97.

Kerr, W. and J. Hobbs. “The North American-European Union Dispute Over Beef Produced Using Growth Hormones: A Major Test for the New International Trade Regime.” *The World Economy* 25, no. 2(2002): 283-296.

Lapan, H., and G. Moschini. “GMO Labeling and Trade: Consumer Protection or (Just) Protectionism?” Paper presented at the 2001 *IATRC meeting*, Tucson, AZ, December 2001.

Lin, W. “Biotechnology: U.S. Grain Handlers Look Ahead.” *Agricultural Outlook*, U.S. Department of Agriculture, ERS, (2000): 29-34.

Macfarlane, R. “Integrating the consumer interest in food safety: the role of science and other factors.” *Food Policy* 27(2002): 65-80.

Malla, S and R. Gray. “Public Research Policy for Today's Agricultural Biotech Research Industry.” *Canadian Journal of Agricultural Economics* 51, no. 3(2003).

Moschini, G., and H. Lapan. “Intellectual Property Rights and the Welfare Effects of Agricultural R&D.” *American Journal of Agricultural Economics* 79(1997): 1229-42.

Pinstrup-Anderson, P. and M. Cohen. 2001. “Rich and Poor Country Perspectives on Biotechnology”, in P. Pardey (ed.), *The Future of Food: Biotechnology Markets and Policies in an International Setting*, IFPRI, Washington.

Qaim, M., and D Zilberman. “Yield Effects of Genetically Modified Crops in

Developing Countries.” *Science* 299(2003): 900-902.

Skerritt, J. 2001. “Biotechnology Policies for Asia: Current Activities and Future Options”, in P. Pardey (ed.), *The Future of Food: Biotechnology Markets and Policies in an International Setting*, IFPRI, Washington.

The Western Producer, April 7, 2005. Monsanto sign royalty deals in S. America, p. 42.

Traxler, G. Personal communication, March, 2005.

Wilson, W., E. Janzen and B. Dahl. “Issues in Development and Adoption of Genetically Modified (GM) Wheats.” *Agbioforum* 6, no. 3(2003): 101-112.

World Trade Organisation, www.WTO.org.

Wright, B.D. 2000. “Intellectual Property Rights Challenges and International Research Collaborations in Agricultural Biotechnology”. In M. Qaim, A.F. Krattiger and J. von Braun (ed.), *Agricultural Biotechnology in Developing Countries: Towards Optimizing the Benefits for the Poor*, Dordrecht: Kluwer Academic Publishers, p. 289-341.

Yampoin, R. and W. Kerr. “Can Trade Measures Induce Compliance with TRIPS?” *Journal of the Asia Pacific Economy* 3(1998): 165-182.

ESSAY 2: A NEW PERSPECTIVE ON AID:

OPTIMAL CONTROL OF EMERGENCY FOOD AID

2.1. Introduction

Food aid has the potential to provide many benefits to recipient countries, especially in emergency food shortage situations. It has been argued that food aid also has the potential to do great harm to recipient-countries' agricultural industries; there exists a considerable literature on the potential negative effects that food aid can have on domestic farm output. The existence of this literature is a positive development, for it shows that there is interest in ensuring that food aid shipments do not just serve the interests of the donor country. That aid be labelled self-serving may sound oxymoronic, since aid, by definition, is intended to help its recipient. The experience with food aid, however, suggests otherwise; donor-country interests were the primary motivation behind the earliest country-to-country aid shipments.

There has been extensive analysis of food aid's effects on the production incentives of recipient-country farmers since Schultz's seminal paper of 1960. This research has been very constructive, as it has shed light on the possible negative effects of food aid.

However there are two key pieces that are missing from the current state of food aid analysis. First, there has not been an attempt to define and identify the quantity of food aid that would be most beneficial for a recipient country. There exists near-unanimous acknowledgement that food aid can impede agricultural production in a recipient country, but there has not been an attempt to identify an amount of food aid that would minimise this damage. Second, the dynamics of existing models are unsatisfying. Those models that examine food aid in dynamic settings introduce food aid as an exogenous shock and trace impulse responses that result from that shock. This technique overlooks the impact that the initial negative supply shock to domestic agriculture (that creates the need for food aid) may have on domestic food production. Also, the introduction of aid may create market conditions in later periods that perpetuate domestic food shortages, thereby creating sustained need for food aid.

This paper formulates a concept of “needed aid” and incorporates this definition of need into a dynamic optimisation model. The model is solved for the optimal path of food aid, and simulations that compare the production and price effects of various aid paths are developed. The objective of this research can be summed up as follows: to construct a model that identifies the most appropriate, or “needed”, quantity of food aid in an emergency situation, and to compare the effects of delivering various amounts of aid on the recipient country’s agricultural industry. The paper proceeds as follows: chapter 2.2 provides a brief overview of food aid, and of the chief donors and recipients. Chapter 2.3 reviews the food aid literature that is relevant to the current analysis. Chapter 2.4 introduces a concept of “need” to the analysis and demonstrates how food aid can be viewed from a new perspective. Chapter 2.5 develops the formal theoretical model that incorporates a definition of need. Chapter 2.6 solves the optimal control model that is developed in chapter 2.5 and simulates various aid paths using generated data. This section also includes a discussion of some policy implications that can be drawn from the model and its simulations. Chapter 2.7 concludes with a recap of the model and its results, and suggests some avenues for further research.

It is worth emphasising that the goal of this research is to model the short-run microeconomic responses of farmers' planting decisions to changes in current and expected market conditions. The goal is *not* to model the economics of famines, *a la* Sen. The forthcoming model assumes that food aid shipments reach those who experience a food shortage. Though such an assumption abstracts from Sen's "entitlement" concerns and from the institutional failures that may exacerbate a famine, it allows the model's focus to remain on food producers' behaviour.

2.2. Primer on Food Aid

This chapter introduces terminology that is common in food aid literature and outlines some of the major trends from recent years. The information in this chapter serves to provide a background for the forthcoming model and to establish some motivation for its creation. There exist several more thorough surveys of food aid trends and policies (see for example, Ruttan or Singer, Wood and Jennings) which should be consulted if a more detailed survey of food aid is required.

The introduction of US Public Law 480 (PL 480) in 1954 marked the beginning of large-scale government-funded food aid programmes. The incentives for PL 480 were two-fold; first to dispose of US farm surpluses that grew out of agricultural support programmes, and second to develop export markets for US agricultural products. PL 480 aid was divided into three categories. Title I aid consisted of US surplus sold to "friendly" recipient nation governments at concessional prices. Title II aid was donated to countries experiencing emergencies, and served primarily humanitarian purposes. Title III aid was donated through non-profit agencies and bartered for strategic materials from recipient countries. A large portion of PL 480 food aid was delivered primarily for the benefit of the donor, not the recipient country (a point which will be addressed in the proceeding literature review). Political and domestic agricultural pressures were often the chief determinants of aid shipments (see the following chapter for a discussion of these motives).

The US was the first nation to institute large-scale government funded aid programmes and remains the largest donor. Ninety-four percent of world food aid programmes in 1965 were comprised of US donations (Maxwell and Singer); this share has fallen over the past forty years, but the US remains the largest donor, currently accounting for 53 percent (Food and Agriculture Organisation). The EU is the second largest donor, accounting for nearly 5 percent and Canada currently accounts for less than 2 percent, down from a peak of 11 percent in 1985. Figure 2.1 shows a brief history of food aid shipments by donor country.

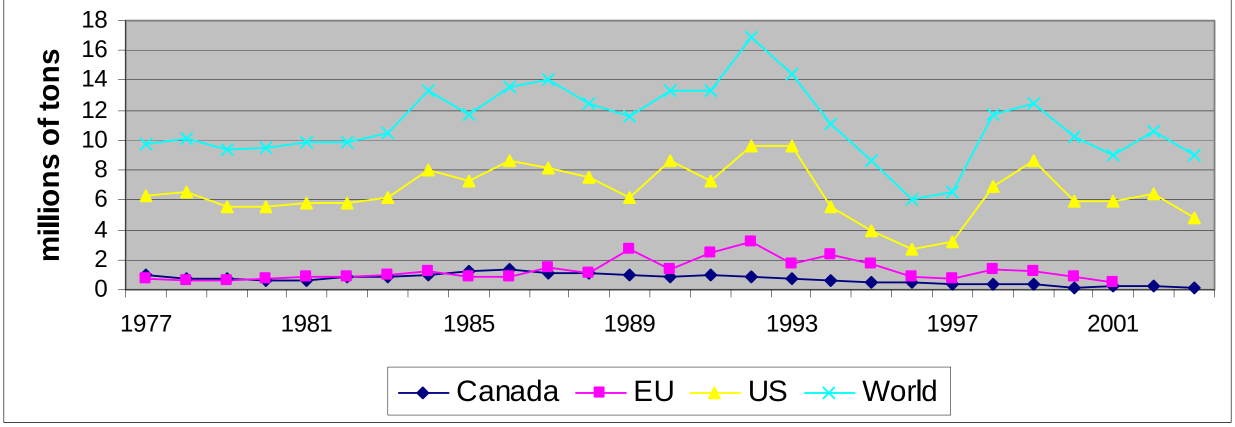


Figure 2.1 - Aid Shipments by Donor (FAO)

Africa and Asia account for the lion's share of the market for food aid shipments, with Central and South America and Eastern Europe comprising most of the remainder. Ethiopia is the largest single recipient, accounting for nineteen percent of total shipments in 2003. Other major recipients in 2003 included Iraq (comprised entirely of emergency food aid), Democratic People's Republic of Korea and Bangladesh (WFP).

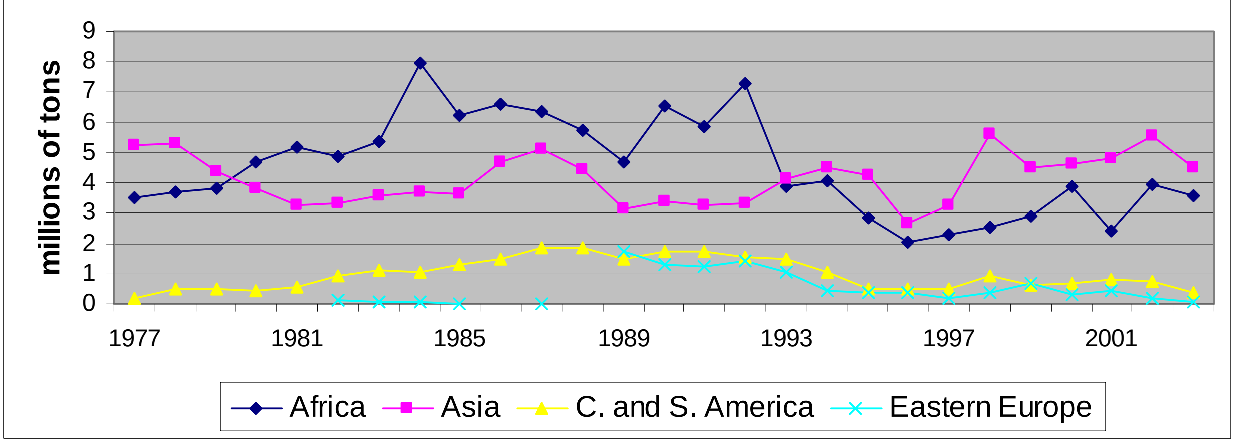
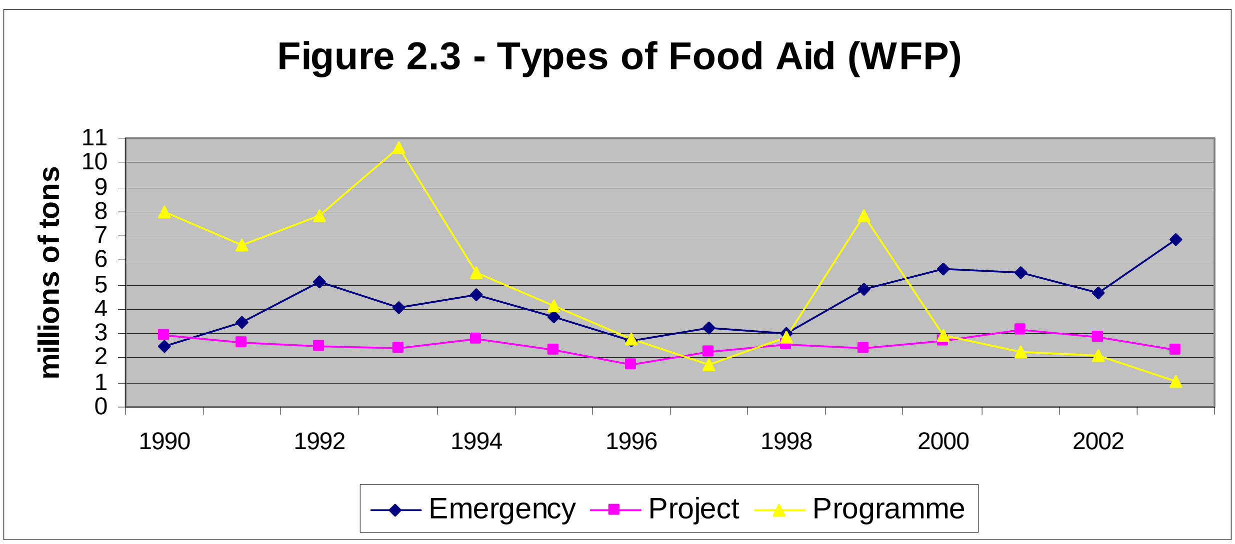


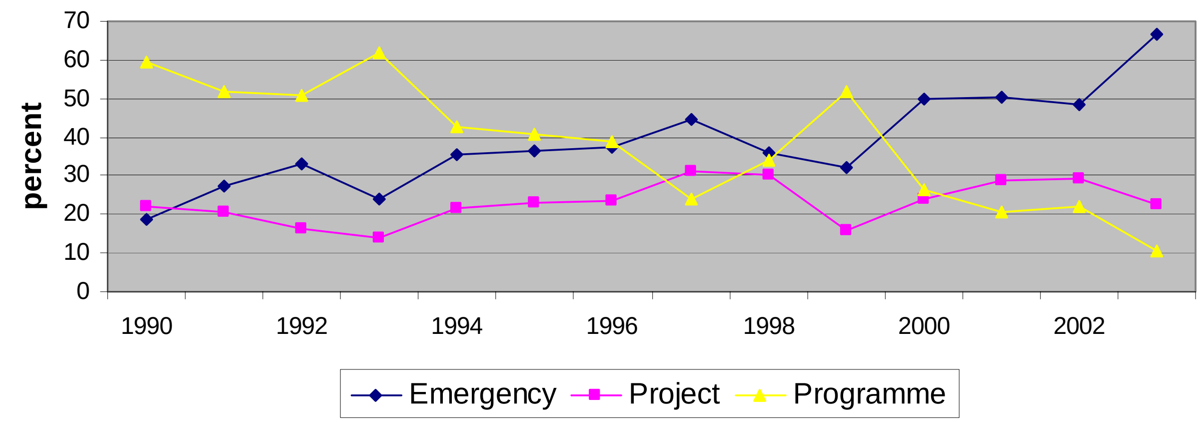
Figure 2.2 - Aid Shipments by Recipient (FAO)

Modern food aid is broadly categorised as one of three types. The World Food Programme (WFP) separates total donations into programme aid, project aid and emergency aid. Programme aid is comprised primarily of nation-to-nation shipments that the donor provides as a means of easing recipient government budget constraints. A recipient government buys programme aid from a donor at a concessional price, and then sells the food in its domestic market. Any margin that the recipient government receives from the sale becomes part of the government's general revenues. Project aid is provided in a similar manner, however the funds acquired from selling the food are earmarked for specific development projects. Most project aid is channelled through multilateral organisations such as the WFP. Emergency aid is provided in times of food crises, often in the event of negative supply shocks or refugee circumstances. Most emergency aid is also channelled through the WFP.



Figures 2.3 and 2.4 illustrate that emergency food aid has evolved to account for the largest share of food aid donations. This increase is the result of rising levels of emergency aid and of falling levels of programme aid. Project aid has remained steady, between twenty and thirty percent of food aid deliveries.

Figure 2.4 - Emergency Aid's Rising Share (WFP)



Most theoretical and empirical food aid studies have addressed the impacts of programme aid (see following literature review). The surfacing of emergency aid as the primary category of assistance since 2000 provides strong incentive to analyse the effects of emergency aid on recipient countries. This is the goal of the current research.

2.3. Literature Review

There exists a large literature on food aid, from case studies to theoretical expositions and empirical analyses. A review of some of the key insights and results from this literature follows.

The first formal analysis of food aid was conducted from the perspective of donor countries. Specifically, Schultz estimated the costs and benefits to the US of surplus disposal under PL 480 during the 1950s. Such analysis was popular at the time because US policy makers wanted to make sure that they received maximum value in return for their food aid shipments. Food aid was primarily a domestic policy tool; the benefits of food aid to recipient countries had not yet come to the fore in policy debate. This attitude is exemplified in a quote from Harold D. Cooly, chairman of the Committee on Agriculture of the US House of Representatives, who stated “We are primarily interested in getting rid of these surpluses and we don’t care how you do it and under what authority. We have told you we want the commodities sold for dollars first and then for foreign currencies or then donate them.”(Schultz)

Several formal attempts have been made to determine the motivation for food aid deliveries. Since a primary motivation for food aid deliveries has historically been to dispose of surplus agricultural production in developed countries, researchers have sought to establish this link empirically, as well as links with other motivations.

Zahariadis, Travis and Ward conduct a two-stage analysis to investigate whether US food aid is a function of political motives or of philanthropic motives. Their analysis uses a two-stage model; stage one is a probit model that estimates the probability of a state receiving aid, and stage two is a multivariate regression that estimates the level of assistance. Zahariadis, Travis and Ward use variables such as domestic cereal production, number of refugees and a proxy for the recipient country’s financial standing as measures of domestic need. Political motivators are measured by reliance on the US for trade in goods and services and by military aid shipments.

Zahariadis, Travis and Ward find that allocation of Title I aid is based on both political and philanthropic motives. Title II (primarily emergency) aid is found to be less statistically dependent on political and philanthropic motives. Title I aid appears to be more politically motivated than Title II, however the statistical relationship between recipient need and Title II aid is weak. Zahariadis, Travis and Ward conclude that since Title II's share of total US aid is rising, political motivators are becoming less important in determining aid flows.

Diven analyses the relationship between donor-country agricultural interests and aid shipments. US aid shipments are modeled as a function of donor stocks, donor exports, lagged aid shipments and recipient country grain production. There are three key results in Diven's analysis. The first is that aid shipments are strongly positively correlated with donor stocks. That is, US aid shipments rise in years of high carryover stocks. Second, US aid shipments are incremental; a given period's aid delivery is highly dependent on the previous period's shipments. Finally, aid shipments are shown to be *positively* related to grain production in the recipient country. Such a finding runs contrary to philanthropic motives. Diven concludes that "US food aid flows have consistently served the interests of [donor-country] commodity producers."

A second focus of food aid analysis is the effect of aid on recipient countries. There are three primary channels through which aid can affect the agricultural industry of the recipient country.

The first is through classical Schultzian disincentive effects. Schultz first addressed this concern in his seminal article that assessed the implications of PL 480 aid shipments. Aid shipments into a recipient country act as an instantaneous supply shock, and have a negative impact on domestic price. Domestic producers then lower production, according to their supply elasticities. This argument can apply to aid that is freely donated or to aid that is sold at a discounted price (providing that arbitrage between recipients of free or discounted aid and consumers of domestic food cannot be avoided).

One of the most contested points in this traditional line of analysis is the size of supply elasticities among agricultural producers in developing countries. Schultz contends that elasticities are significantly different from zero (contrary to what Schultz describes as "the widely held belief that the price response of cultivators is zero"). Likewise, Fisher argues that a supply elasticity of zero implies that agricultural inputs have no alternative use. This seems implausible, especially as economies develop and achieve faster growth. Rogers, Srivastava and Heady concur with this view, and suggest that "the proposition that production in developing countries is not price responsive has little basis." It seems rational that agricultural producers in developing countries are price responsive.

The analysis of food aid's disincentive effects can also be extended to consider long-term consequences. Two long-term effects are particularly worthy of note. First, if there is a short-term decline in agricultural production, then physical capital may depreciate and labour may relocate away from farms. If this capital and labour is not replaced, then long-term production may not recover. Also, food aid may take the pressure off of recipient-country governments to invest in, and pursue, policies to develop domestic agricultural production (Rothschild). Politicians may feel that food aid provides an opportunity to focus policies on urban and industrial development at the expense of domestic agriculture.

A second channel through which aid can affect recipient country agricultural markets is by displacing commercial food imports. A recipient country that finds itself with free or discounted food shipments may consequently reduce its demand for imported food. Some donors impose Usual Marketing Requirements (UMRs), which stipulate that recipient countries cannot reduce commercial imports of agricultural products from donor countries during periods of aid shipments - food aid must be "additional" to current imports. UMRs are often disobeyed, and even if they are observed, the recipient country may simply reduce demand from other exporting countries.

A third possibility is for food aid to have a positive effect on the development of a recipient country's agricultural markets. Food aid, if additional consumption for recipient country workers, could have positive nutritional and health effects (Barrett, Mohapatra and Snyder). Such effects could increase the productivity of labour, thereby increasing agricultural output. Similarly, if food aid displaces commercial imports and frees up foreign exchange, donor countries could spend this currency on imported capital to increase agricultural productivity. Rogers, Srivastava and Heady suggest that food aid may trigger an income effect that pushes out aggregate demand for food and exerts upward pressure on local food prices. Such an effect could help to offset the disincentive effects of increased supply.

There have been several attempts, both theoretical and empirical, to formally analyse the effects of food aid on recipient-country agricultural markets. These analyses can be broadly grouped into two categories. The first category includes older models that employ comparative statics to trace the effects of aid on recipient-country food production. The second, more contemporary, category includes time series econometric models that include aid shipments as an endogenous variable in a series of estimated equations. A brief review of some important findings from each category follows.

Fisher was one of the first to present a formal theoretical analysis of the impact of foreign surplus disposal on recipient-country agriculture. Fisher’s model formulates a standard demand and supply comparative static framework that interprets the effects on domestic supply through a supply elasticity. The effect is illustrated by a rightward shift of the supply curve in the recipient country and a resulting lower price offered to domestic producers. The magnitude of the decrease in local production is determined by the price elasticity of supply.

Two important points can be drawn from Fisher's work. First is that there existed a need for empirical estimates of supply and demand elasticities in developing countries to implement Fisher's model. Demand elasticities were required because Fisher's work suggests that lower food prices may increase consumers' real income, thereby increasing demand for food. This proposition is similar to that put forward by Rogers, Srivastava and Heady.

Fisher also stresses the importance of estimating how large policy expenditure must be (in the form of subsidies to farmers) to offset negative price effects of imported food aid. Such funding could come from government general funds or from the proceeds of selling Titles I and III aid.

A second important theoretical analysis of the effects of food aid is Srinivasan's production possibility frontier (PPF) model, within the context of food aid's general equilibrium effects on Indian development. Food aid unambiguously increases aggregate welfare in the two-good (food and non-food) world of PPF analysis, though the relative price of food falls. Srinivasan suggests that since the economy is better off, the welfare gain can be redistributed to compensate food producers, whose terms of trade have fallen. It seems that Srinivasan is suggesting that food aid passes a Hicks-Kaldor compensation test.

Several authors have attempted to quantify the effects of food aid by means of estimating elasticities. Mann was one of the first to estimate a system of structural equations that identifies a relationship between aid shipments, commercial imports, domestic food price and domestic food production. Mann's model utilises a static expectations formulation to describe producers' response to price changes. His results show that aid deliveries have a concurrent depressing effect on domestic price and a negative two-period lagged effect on domestic output. The two period delay in supply response is the result of the lag structure in Mann's supply equation; output in the current period is a function of planting in the previous period, which is a function of price two periods prior.

Mann's simulation results exhibit almost hysterisis-like patterns. A one-time aid shipment shock causes price to fall in period zero, and planting to fall in period one. Production consequently declines in period two, resulting in a decline in market supply of food. This pushes domestic price up and increases planting one period hence. Domestic production cycles around its original level every two periods, with decayed effects in each cycle. Mann's analysis identifies significant disincentive effects of food aid on domestic agricultural production.

Rogers, Srivastava and Heady take a similar approach to Mann, but consider the effects of segmenting the domestic food market into two pieces; one for sales of domestic production and another for concessional sales of imported food aid. Rogers, Srivastava and Heady arrive at similar conclusions to Mann, however they find that disincentive price effects can be reduced if food aid is sold at a discounted price in a concessional market. This conclusion is, however, contingent on no arbitrage between the two markets.

The last few years have witnessed a change in the direction of food aid analysis. Modern time-series econometrics have provided researchers with new techniques for analysing the effects of food aid on recipient-country food production. Vector autoregressions (VARs) have become the models of choice. VARs are popular for several reasons. First, the data requirements are less than for structural models. Food production need not be modeled as a structural supply function in a VAR; rather it is modeled as an endogenous variable in a system with food aid shipments, price and commercial imports. VARs are also popular because they can appropriately model nonstationary data, which would otherwise lead to spurious regression results (Granger and Newbold).

Donovan, et al. use VAR estimation to simulate the effects of food aid on maize prices in Mozambique. Their model estimates a system of VAR equations with food aid shipments and prices of white and yellow maize as endogenous variables. Simulations show that maize prices would have been higher in Mozambique in the absence of food aid shipments. Donovan, et al. conclude that aid shipments to Mozambique have benefited consumers by means of lower food prices. However, the flip side to this benefit is that “a disincentive effect on domestic production and marketing cannot be ruled out in the longer run.”

Barrett, Mohapatra and Snyder follow a similar approach, but include domestic production, rather than domestic food price, as an endogenous variable in a VAR analysis. Barrett, Mohapatra and Snyder's goal is to determine if the data reveal a negative correlation between food aid shipments and commercial imports or a negative correlation between food aid shipments and domestic food production. Barrett, Mohapatra and Snyder identify a short-run fall in domestic production in response to aid shipments, which they attribute to price disincentives. Food aid also displaces commercial imports in the short run, thereby providing the recipient country with a de facto foreign exchange transfer. Domestic food production recovers over the long term (which the authors define as twenty years); this recovery is attributed to the benefits of the recipient country's ability to import new physical capital, using the foreign exchange that would otherwise have been spent on commercial food imports. Barrett, Mohapatra and Snyder's conclusion is that "the data support the Schultzian critique that food aid discourages recipient country production in the short run."

Note that most of the aforementioned studies analyse the effects of PL 480 Titles I and III (non-emergency) food aid. Such aid is persistent (Barrett, Mohapatra and Snyder) and often motivated by factors other than recipient need (Diven), (Zahariadis, Travis and Ward). One key exception is Donovan, et al.'s analysis. Though Donovan, et al. do not detail the source and category of the aid shipments used in their estimation, it appears as though they focus on emergency aid.

Despite seemingly inconsistent results from the vast landscape of food aid analyses, a few broad conclusions permeate most studies. First, the data support the theoretical notion of a Schultzian disincentive effect in the short run. Dynamic time-series models tend to show that the disincentive effect decays over the long term, and domestic production recovers. Second, most models demonstrate a negative relationship between food aid shipments and commercial food imports, thus violating the additionality principle[28]. It seems that the debate is not whether food aid affects the recipient country agriculture market. Rather, the debate is how large is the effect and how long it lasts.

2.4. A New Perspective

The idea for this research grew out of discussion about WTO trade rules governing export subsidies. Specifically, how can the WTO institute a set of rule-based disciplines that prevents member countries from using food aid to skirt the spirit of the WTO Agreement on Agriculture, without jeopardising legitimately-required food aid? The WTO currently defers to the FAO for rules that determine the legality of food aid. These rules determine legitimacy of food aid from the donor's perspective; a proposition that is wrought with difficulties. Crafty lawyers in donor countries seem always able to invent channels through which food aid can be deemed "WTO legal". For example, the FAO *Principles of Surplus Disposal and Consultative Obligations of Member Nations* (FAO 1992) distinguishes food aid from commercial food trade by measuring the concessionality of the food transfer. As Shaw and Singer point out, the meaning of "normal commercial practice", which is key to the commercial food trade

[28] The additionality principle states that food aid should be wholly additional to, rather than a substitute for, commercial food imports.

definition, is different among governments. “Normal” is a vague and subjective term that makes the implementation of a consistent set of rules difficult. It would seem that another method of determining the “legitimacy” of food aid would be beneficial.

Since the primary goal of food aid should be to benefit recipient countries, it is logical that the legitimacy of food aid should be determined from the perspective of recipients. That is, food aid should be regarded as legitimate if it benefits the recipient, and illegitimate if it does not. How does one determine if food aid benefits a recipient country? For this we focus on the case of emergency food aid, and try to formulate a concept of “need”.

2.4.1. Defining Need

“Need” is a vague term that is difficult to define in the context of development economics. On one hand, developing countries need a quantity of food to meet nutritional requirements. Does a developing country, however, need to produce this food self sufficiently (as opposed to importing)? Perhaps not, but few countries have achieved modern economic growth without developing a productive agricultural industry (FAO, 2001). We proceed under the notion that domestic agricultural production is beneficial to developing countries. Also, to limit the scope of a definition of need, we focus on emergency food aid. How much food aid does a recipient agrarian country need when faced with a negative supply shock? There are several *a priori* responses to this dilemma:

i. As much as possible. As economists, we are taught that consumers are insatiable, and that disposal is free. This line of reasoning would argue that there could be no such thing as too much aid.

ii. No aid. This policy states that the provision of any food aid only disrupts the development of market institutions that could help to alleviate the negative effects of a future supply shock. Such institutions could include storage facilities and crop insurance programmes. The "no aid" policy is appealing from a moral hazard perspective. Moral hazard in the context of food aid suggests that by granting aid to developing countries, recipients are more likely to put themselves in situations that require further aid in the future. Such a concern could be alleviated by providing no aid.

iii. The exact amount of food lost to a negative supply shock. It makes sense, *a priori*, that just replacing the amount of lost food would have the smallest effect on the recipient-country's agricultural markets; as if the shock didn't happen.

I argue that none of the aforementioned approaches accurately identifies the recipient country's need in times of emergency. The first option is sure to produce Schultzian disincentive effects. Most authors agree that there exists a real risk of depressing local food prices with food aid, even if none can agree on the magnitude of this risk.

The second response is unappealing for humanitarian reasons. Acute food shortages must be addressed with food aid, lest people starve. Non-interference may be an appealing policy in cases of chronic, mild shortages, but the need to keep people alive in emergencies must trump all other concerns. Isenman and Singer extend a similar hands-off line of reasoning. They argue that price signals that result from shocks, and potential disincentive effects that are created by aid policies, may not be such bad things if they encourage farmers to switch to different, perhaps more profitable, crops. However, emergencies are not the appropriate occasions to make such switches; long-term price signals and incentives should determine resource allocation. Also, the moral hazard dilemma is not addressed in the current research. The large-scale macroeconomic and institutional developments required to prevent future food emergencies are outside the scope of this research.

The third response is appealing in that it avoids nutritional shortfalls and should avoid Schultzian disincentive effects. Market supply of food does not increase if food aid just replaces the amount lost to a shock, so there should be no downward pressure on food prices. However, there may be factors other than current price that affect producers' ability and desire to change production levels. Such factors may dictate alternative aid levels, depending on the definition of need. This point is addressed in detail in the forthcoming model and simulations.

A working definition of need is required to proceed with a model. The first requirement of this definition is that people do not starve; food aid shipments must at least make up for the effects of a supply shock. This is referred to as the "nobody starves" condition. The second requirement is based on the assumption that agricultural development is a worthwhile endeavour in the recipient country, and that an emergency situation is not the ideal opportunity to apply pressure for structural or institutional change. As such, aid shipments should be allocated in a manner so as to minimise negative effects on local agricultural markets. Note that the "nobody starves" condition is not the same as a notion of long-term food security. Food security is a difficult concept to measure (Gray); for example a food aid policy may seek to maximise growth subject to a minimum level of food security (Gray). Emergency aid is not well-suited to such a policy objective. The primary goal of emergency food aid should be to avert starvation and acute malnutrition, and to allow a recipient country's agricultural industry to recover as quickly as possible from a supply shock. Increasing the nourishment of residents beyond a pre-shock level and increasing productivity in a recipient country's agricultural industry are goals that should be targeting by other policies (such as institutional reform and programme aid). A working definition of need in the case of emergency food aid is then:

> Emergency food aid is needed in an amount that minimises potential damage to the recipient country's domestic agricultural industry, subject

to at least making up for the nutritional shortfall imposed by an emergency.

The “nobody starves” condition appears in the definition as a constraint, i.e. “subject to at least making up for the nutritional shortfall…” It is important to note that meeting this restriction should be the fundamental goal of any food aid policy. Including the “nobody starves” condition as a constraint rather than an objective is simply a matter modelling semantics. This is point is addressed further in chapter 2.5, which outlines the technical model.

2.4.2. Modelling Need

The problem now is to apply this definition of need to a model that simulates the effects of different levels of aid on a recipient-country’s agricultural industry, and determine what level of aid produces the most beneficial (from the recipient country’s perspective) results. A description of the elements required for a food aid model follows.

A dynamic structure is important in a food aid model. The agricultural production process, and its response to supply shocks and food aid shocks, is inherently dynamic. As such, any model that seeks to characterise how food aid affects a local agricultural market must consider how decisions in the current period affect output, price and aid deliveries in the next period. The size of a supply shock and subsequent aid deliveries both affect production decisions, thereby impacting the need for aid one period hence. Static models overlook these important inter-period linkages and are unsatisfying for two primary reasons. First, they rely on price elasticities, which may not consider the effects of the initial supply shock on productive capacity. Second, the lag between planting and harvest is not considered in static models.

Dynamic time-series models are also inadequate for analysing such dynamics, especially from a recipient-country perspective. Food aid is commonly analysed in VAR models by imposing an exogenous food aid shock and tracing impulse responses (Donovan, et al., Lowder). Food aid may continue past its initial arrival, but not according to need, as evaluated in each period. Rather, aid's impulse response is a function of estimated VAR parameters.

Since the objective of the current research is to identify the most appropriate, or needed, quantity of food aid, we require a policy variable. The logical choice for this variable is food aid shipments. Aid shipments are assumed to be determined by policy makers from the local government, the WFP or bilateral donors. The model contains a mechanism by which aid shipments affect local price through its effect on domestic food supply.

The food aid model also requires a market supply response. Market supply is made up of food aid and domestic production. The responses of domestic farmers to three factors must be considered. First, the initial supply shock that initiates an emergency.

A supply shock in the form of a crop failure affects revenues, thereby decreasing farmers' abilities to hire inputs, pay wages and finance other production costs. This factor has heretofore been unaddressed in food aid models, but is as important a factor in supply responses as potential price disincentive effects.

Price effects are the second factors to consider. Price affects production through two avenues. First, current price affects current revenue, which impacts production as previously described. Second, farmers base current planting decisions on their expectations of price in the harvest period. As such, expected future price affects current planting.

Finally, the initial supply shock and subsequent production decisions may affect the speed at which a farmer can vary his level of output. Specifically, a large farm with many inputs could decrease output more rapidly than a small farm with few inputs could increase output. Decreasing output merely requires the disposal or idling of inputs, while increasing output may require the acquisition of physical capital, labour and skills; all of which may take several periods to obtain. The specifics of these factors are addressed in chapter 2.5, where the technical model is fully explained.

The pieces of the model just described fit well into a dynamic optimisation framework. There is a policy, or control, variable (food aid shipments), a behavioural or state equation and variable (supply response and food supply) and a relationship between the two (price determination). Such a framework shares much in common with natural resource models that identify optimal harvest rates of renewable resources. A common natural resource corollary (Leonard and Van Long) is the fisheries example. Fishing effort is the control variable, the natural growth rate of the fish stock is the state equation and the effect of harvest on fish stock is the relationship between the two.

One more piece is required to model food aid in a dynamic optimisation framework; an objective function to optimise. Based on the previously defined notion of need, we try to operationalise an objective function that embodies the need to replace the food lost to a supply shock and the desire to avoid damage to the recipient country's agricultural development. Prime candidates for defining this damage are food price volatility and the difference in domestic output from its pre-shock level.

2.5. Model

The model follows dynamic optimisation conventions by defining a control variable, a state variable, a state equation and an objective function. The goal is to identify the optimal control path of the policy variable that minimises the stated objective function.

The control variable is food aid shipments. Food aid deliveries are determined exogenously by a planner and are not subject to the exogenous supply shock that affects domestic harvest. The planner can be a government agency in the recipient country, a foreign multinational organisation such as the WFP or a single-donor country. The key to the control variable is that it is substitutable for domestically produced food, and is not purchased on the local market. It is imported from abroad and is wholly additional to domestic food supply. The manner in which food aid is distributed is discussed further in the following chapter that describes the price determination equation.

Some models (Rogers, Srivastava and Heady) propose that food aid should be considered in a two market model; one market for domestic supply and a separate market for aid that is either freely distributed or sold at concessional prices. Such models are applied to programme aid, and are less suited to emergency situations. It seems unlikely that a planner could enforce non-arbitrage between two such markets in emergency situations. Enforcement would be particularly difficult in emergency conditions because arbitrage opportunities are high and government and aid resources are in great demand. The control variable, then, is defined as y_t; aid shipments in volume in period t (for example, tonnes per year).

We now turn to the state variable and its equation of motion. As discussed in the previous chapter, the state variable is domestic crop planting. Domestic food supply is modelled to respond to market conditions, and is characterised by a representative agent farmer. All farmers are assumed to respond identically to a supply shock and to price signals. Current period crop planting is x_t. Note that current planting, x_t, represents the quantity of seeds planted in period t (for example, tonnes of seed per year), not the area of seeded land.

Note that x_t does not represent current food supply. Due to the production lag between planting and harvest, we represent current harvest independently of x_t. Current period harvest is h_t; $x_{t-1} = h_t$ unless a supply shock disrupts harvest. It is important to remember that x_t, and not h_t, is the state variable. Current period harvest is predetermined in period t, and only differs from x_{t-1} by the amount that an exogenous supply shock decreases harvest.

Before outlining the state equation, we must make some background assumptions about the recipient country's agriculture industry. First, the recipient country's agriculture industry is in long-run equilibrium wherein production and prices are in a steady state. For this to be the case, we must make a second assumption - that demand for food is fixed. This level of demand, hereafter referred to as A, could represent the minimum sufficient caloric intake, as determined by medical science. For simplicity, we further assume that the recipient country's steady-state output is equal to A, so that the country is self-sufficient in food production without food aid. This assumption is not essential to the model - the model could be specified so that the steady-state equilibrium output is equal to A less food imports or aid - but doing so simplifies the algebra and clarifies solution insight.

Autarky is another of the model's assumptions. This assumption allows the model's focus to remain on recipient-country producer behaviour. Also, an autarkic situation allows price to be determined locally, so that food prices are not fixed at the world level. Though autarky is one end of the trade openness spectrum, it is not an unreasonable assumption to make in the case of several countries that are affected by emergency food supply conditions. Ethiopia, for example, ranks low on measures of economic openness (Kandiero and Chitiga), and is a frequent recipient of emergency aid (WFP).

The primary drawback of analysing food aid in the context of an autarkic nation is that we overlook the possible important effects that aid may have on commercial imports. Several studies (Barrett, Mohapatra and Snyder, and Bezuneh, Deaton and Zuhair) have analysed the possible displacement effects that food aid might have on commercial food imports, and these results are insightful. However the primary goal of this paper is to focus on the recipient-country producer response, and as such we proceed with an autarkic model.

Given the assumption of autarky, we further assume that domestic food price is determined through an excess supply function (Varian). Given fixed and constant domestic demand, market price fluctuates as domestic food supply changes relative to domestic food demand. Food price begins in equilibrium, where supply equals demand. If food supply rises above food demand, then price falls below its pre-shock level. Such a relationship can be represented by

$$P_t = \delta\left(\frac{A}{h_t + y_t}\right) \tag{2.1}$$

Current period price is determined by the ratio of food demand to food supply, through a positive parameter δ. Food aid is assumed to be freely distributed among residents of the recipient country and arbitrage between domestic food and imported aid cannot be avoided. Arbitrage between free and purchased food will provide an average market food price that will be faced by local agricultural producers.

The state equation is comprised of three parts, each of which impacts local farmer production decisions, thereby altering the state variable (current period planting). The first component of the state equation embodies the immediate effect of the supply shock on the farmer's ability to pay input costs. We assume that a farmer forms an expectation of P_t in period $(t-1)$, which is referred to as $E_{t-1}(P_t)$, and therefore expects to receive revenue equal to $h_t E_{t-1}(P_t)$ in period t. If the farmer's price expectation is correct and there is no supply shock (so that $h_t = x_{t-1}$) then

$h_t P_t = x_{t-1} E_{t-1}(P_t)$ and the farmer receives his expected revenue. Output remains at its pre-shock level and the state variable does not change. To account for the possibility of an exogenous supply shock, the relationship between planting in period $(t-1)$ and farmers' expected harvest in period t is represented as $x_{t-1} = h_t \left(\frac{1}{1-\Omega_t} \right)$. Ω_t represents the size of an exogenous supply shock in period t; for example, a 75 percent negative supply shock results in $\Omega_t = 0.75$. In the case of no supply shock, $\Omega_t = 0$ and $x_{t-1} = h_t$. Farmers' expected revenue in period t is represented by $h_t \left(\frac{1}{1-\Omega_t} \right) E_{t-1}(P_t)$. Returning to the case of a 75 percent negative supply shock in period t, a farmer who planted 10 units in period $(t-1)$ would harvest only 2.5 units in period t and would receive revenue of $h_t P_t = 2.5 P_t$. This compares to expected revenue of $2.5\left(\frac{1}{1-0.75} \right) E_{t-1}(P_t) = 10 E_{t-1}(P_t)$. The expected revenue term can be simplified as $h_t \pi_t E_{t-1}(P_t)$, where $\pi_t = \left(\frac{1}{1-\Omega_t} \right)$. Note that π_t will equal 1 in all periods in which an exogenous supply shock does not occur.

If, however, the farmer's revenue expectations are not correct, then there is pressure for output to change. Consider the case where a negative supply shock reduces harvest in period t, so that $h_t < x_{t-1}$ and price does not change. If the farmer conducted forward contracts with his inputs in period $(t-1)$ with the expectation of paying these inputs out of revenue generated in period t, then the farmer experiences a revenue shortfall; $h_t P_t$

is less than $h_t \pi_t E_{t-1}(P_t)$. Such contracts could include an indenture to pay labour out of subsequent harvest or contracts wherein the farmer borrowed to acquire capital with the intention of repayment one period hence. If there is a revenue shortfall, then either these inputs must go unpaid or the payment must come from a source other than current period's revenue. If labour is unpaid, then workers may leave the farm, and if capital loans are not repaid, then the capital may be repossessed. If the funds to pay these inputs are acquired from other sources, then there are two possible sources. One is the stock of funds that was intended to pay for variable inputs in period t to plant x_t (for example, labour or fertilisers). The second source is funds acquired from the sale of currently owned capital, land or other fixed inputs that would have otherwise been used in current period production. Either of these scenarios leads to lower productive capacity and has a negative effect on output in the current period.

This *revenue effect* is operationalised by

$$\left(h_t P_t - h_t \pi_t E_{t-1}(P_t)\right) \tag{2.2}$$

which is negative if actual revenue is less than expected revenue.

The second component in the state equation is the *expected price effect*. This effect is straight forward in that the farmer wants to increase production if he expects the price of his crop to rise above the pre-shock price. Stated explicitly, if $E_t(P_{t+1}) > P_0$, then the farmer seeks to increase planting in the current period (x_t) in hopes of increasing next period's harvest (h_{t+1}) to take advantage of the higher price. Note that $E_t\left(P_{t+1}\right)=P_0$ in the steady state, and planting is constant. However if $E_t\left(P_{t+1}\right)$ deviates from P_0, then the state variable, current period planting, deviates from its pre-shock level.

The expected price effect is operationalised by

$$\left(E_t\left(P_{t+1}\right)-P_0\right). \tag{2.3}$$

A brief side note is required at this point to explain the expected price terms utilised in the first two components of the state equation. The model incorporates farmers' price expectations as part of the production decision and these expectations require specification within the model. There exists no consensus on how price expectations should be modeled in agricultural economics. Empirical tests of various expectation hypotheses in agriculture (Kenyon, and Shideed and White) tend to favour expectations based on either current prices (naïve expectations) or on futures markets. Futures markets do not exist in most emergency food aid recipient countries, and naïve expectations do not seem entirely appropriate to emergency food situations. There is no reason to think that farmers should expect that the price of food in an emergency period should be the same as the price of food before or after an emergency situation. Supply conditions are sure to change in the periods following a supply shock. Fortunately, the structure of the control model and the steady-state conditions provide some direction on how to model price expectations.

The farmer forms his price expectation with the knowledge of fixed demand, A, and therefore bases his price expectation on his projection of food supply in period $(t+1)$. This is akin to equation (2.1), wherein price is determined by an excess supply function. To formulate a belief about supply one period hence, the farmer must have a projection of the amount that he will plant in the current period. This, however, creates a simultaneity problem; price expectation is part of the state equation that determines output, and projected change in output forms the basis of the farmer's price expectation. To get around this problem, we require a proxy for the farmer's output capacity. For that proxy, we use variables that provide some insight into how much the farmer will produce, and that are determined prior to period t's planting decision. A reasonable proxy for the farmer's productive capacity in period t is the farmer's crop revenue in period t. As crop earnings rise, the farmer plans to increase output and expects food supply to rise in period $(t+1)$. We therefore express the farmer's price expectation as

$$E_t\left(P_{t+1}\right)=\frac{\gamma}{h_t P_t}. \tag{2.4}$$

The denominator of this expression, $h_t P_t$, is a proxy for the farmer's current productive capacity and the numerator is a positive parameter. As the farmer's current revenue increases, his productive capacity rises. The farmer's proxy for current period planting (current revenue) rises, and he therefore expects that harvest in period $(t+1)$ will rise above current period harvest. The farmer bases his price expectation on a pseudo-excess supply function, and therefore expects that as supply rises relative to demand, price will fall. Higher current period revenue results in a lower expected price in period $(t+1)$. Note that the parameters in equations (2.1) and (2.4) can be calibrated so that $E_t(P_{t+1}) = P_{t+1}$ in the pre-shock state.

The state equation's third element incorporates the speed at which a farmer can change his level of output. The rationale for including such a factor is that a farmer should be able to decrease production from the initial steady-steady state level more quickly than he can increase production to return to that initial output level.

Emergency situations can have marked effects on input allocation within a region. If farm revenue falls short of expectations, wages to farm labourers may not be paid and labourers may relocate - often to urban areas. Also, if current period planting falls, so too do labour requirements; some of this unhired labour may relocate away from farms. A similar effect can occur with non-labour inputs. Capital or land that is not used in current period production can be idled or sold. The result is a shift in input allocation away from farms. As a means of incorporating this dynamic into the optimal control setting, a farm's capacity to change its level of production is modeled as being proportional to its output one period earlier. Including $h_t\pi_t$ (which is equal to previous period's planting, x_{t-1}) in the state equation allows the model to respond so that a large farm can decrease output by a larger amount than a relatively small farm can increase output.

The state variable is defined as

$$x_t = x_{t-1} + \rho h_t \pi_t \left[\alpha\left(h_t P_t - h_t \pi_t E_{t-1}(P_t)\right) + \beta\left(E_t(P_{t+1}) - P_0\right)\right]. \tag{2.5}$$

Equation (2.5) states that current period planting is equal to previous period planting plus the change in production as defined above. ρ, α and β are positive parameters. Moving x_{t-1} to the left-hand side of the discrete form of equation (2.5) and taking the limit as $\Delta t \to 0$ provides the state equation, or equation of motion for crop planting:

$$\dot{x} = \rho h_t \pi_t \left[\alpha\left(h_t P_t - h_t \pi_t E_{t-1}(P_t)\right) + \beta\left(E_t(P_{t+1}) - P_0\right)\right]. \tag{2.6}$$

It is worth noting that the discussion about farmer supply responses does not include classic Schultzian disincentive effects. Classic Schultzian disincentive effects entail an immediate negative supply response in reaction to falling current food prices; such a response is inappropriate for a dynamic model. Planting decisions are made one period before output is realised in a dynamic model to account for the planting-harvest lag. It seems logical, then, that rational farmers base their planting decisions on their expectation of next period's price, rather than current price. It is, after all, next period's price that farmers receive for currently planted crops. Current planting may decline in response to a fall in current price (which may fall as a result of excess aid), but this decline is attributable to revenue effects, not price disincentive effects. Also, current period planting may decline in response to price disincentives, however the disincentive is next period's price, not current price.

The final element of the optimal control model is the objective, or value, function. The objective function is how the aforementioned definition of "need" is incorporated into the model. The goal of the optimal control model is to determine the path of the control variable, food aid shipments, that minimises damage to the recipient-country's agriculture industry. To determine this path, we require an explicit formulation of damage. Working under the proposition that agricultural production is a worthwhile endeavour in the recipient country, damage is defined in two ways. The first is a measure of price instability, which is calculated using squared food price differences from one period to the next. The second measure of damage is the difference between food requirements (the constant A, as defined above) and current period planting (x_t). x_t is equal to A in the pre-shock state, so the larger is this difference, the more damage has been done to the recipient-country's agricultural industry. Per period damage, or loss, can then be stated as

$$\left[(P_t - P_{t-1})^2 + (A - x_t)\right]. \tag{2.7}$$

The second term in the per-period loss function is not squared because a production capacity constraint is imposed on local production. Specifically, local planting is limited to the constant food requirement so that $x_t \leq A$ and the term $(A - x_t)$ cannot be negative. This constraint corresponds to the initial self-sufficiency of the recipient-country market, and is appropriate for the goals of this model. This model seeks to identify the short-run micro responses to emergency food aid, and not the long-run growth issues that may be associated with programme food aid.

The cumulative loss to the recipient country can be obtained by taking the summation of equation (2.7) over a fixed time period. Total loss is represented as

$$L = \sum_{0}^{T} \left[\left(P_t - P_{t-1} \right)^2 + \left(A - x_t \right) \right] \tag{2.8}$$

where t represents time from periods 0 to T.

Three comments are necessary about equation (2.8). First, there is no mention of the "nobody starves" restriction in the control model's objective function. This restriction is introduced below, as equation (2.12). Recall, however, that the way in which the "nobody starves" condition is modeled is simply a matter of semantics. The primary goal for an emergency food aid policy should be to meet nutritional requirements; including this goal as a restriction instead of an objective is a matter of model technicalities, not an indication that it should be an afterthought.

Second, there is no discount factor, as is often present in the value functions of natural resource models. This loss function is not specified in dollars, so a discount factor is not included. Third, it is the ordinal value of L that is of interest. The units in equation (2.8) do not have intuitive meaning and there are no weighting parameters attached to the loss function's two components; the cardinal value of L is unclear. The comparison of L between different paths of food aid is the relevant focus of evaluation. The specification of the value function is discussed further in chapter 2.6.3.

Now that all components of the optimal control model are defined, it is worthwhile to present all of the model's pieces together:

$$\min_{y_t} \sum_{0}^{T} \left[\left(P_t - P_{t-1} \right)^2 + \left(A - x_t \right) \right] \tag{2.9}$$

subject to

$$\dot{x}(t) = \rho h_t \pi_t \left[\alpha \left(h_t P_t - h_t \pi_t E_{t-1} \left(P_t \right) \right) + \beta \left(E_t \left(P_{t+1} \right) - P_0 \right) \right] \tag{2.6}$$

$$x_t \leq A \tag{2.10}$$

$$x_0 = A \tag{2.11}$$

$$y_t \geq A - h_t \tag{2.12}$$

$$\Omega < 1 \tag{2.13}$$

$$\lambda_T = 0 \tag{2.14}$$

where

$$P_t = \delta\left(\frac{A}{h_t + y_t}\right) \tag{2.1}$$

$$E_t\left(P_{t+1}\right) = \frac{\gamma}{h_t P_t}. \tag{2.4}$$

The only new components to this model are constraints (2.12), (2.13) and (2.14). Inequality (2.12) constrains the control variable, food aid shipments, to be greater than or equal to the area's subsistence requirements less the current period harvest. Put another way, food aid shipments are required to at least make up for the harvest lost to the supply shock that precipitates the emergency. This constraint is included for humanitarian reasons. The preservation of life is not explicitly incorporated into the loss function, but it is assumed that the loss of life to starvation trumps any price and production disincentive effects of food aid. Food aid shipments are therefore restricted to always provide at least as much food as is required in the recipient region (the "nobody starves" condition). The goal of this model is to focus on the needs of the recipient country; a goal that would be grossly unattained if residents in the recipient country were to starve. Inequality (2.13) constrains the size of the supply shock to be less than absolute. This is a realistic constraint, in that a supply shock is not likely to ever wipe out every trace of planted crop. Also, a total supply shock, or $\Omega = 1$, renders π indefinable. Inequality (2.14) is the transversality condition required in an optimal control model with fixed terminal time (Kamien and Schwartz). This requirement states that the costate variable must equal zero in the final time period. In other words, the shadow value of current planting (the intuitive meaning of the costate variable) must not have any influence on the objective function in the terminal period.

The control problem is specified from the perspective of the central aid authority. The aid authority determines a quantity of aid to deliver in each period in an effort to minimise the cumulative loss function of equation (2.9). The aid authority has complete knowledge of the model's parameters, including the specification of farmers' price expectations. The aid authority therefore has perfect foresight of the consequences of its aid policy.

The dynamics in the economy described in the optimal control model unfold as follows. The economy begins in a steady-state equilibrium with productive capacity equal to A. Prior to a supply shock, local farmers plant $x_t = A$ and receive harvest of $h_{t+1} = A$, which can be sold at the steady-state price of P_0. The economy remains in this steady state without a supply shock. Now consider a supply shock that affects the harvest level in period 1; such a shock could be inclement weather, infestation of pests or violence that prevents harvest. This shock breaks the direct link between planting and subsequent harvest, so that $h_1 < x_0 = A$. The relevant food aid authority observes this shortage and responds by delivering food aid in the amount y_1. The market price for food, P_1, is then determined by equation (2.1) and recipient-country farmers form their expectation of P_2 according to equation (2.4). Farmers then make their planting decision, x_1, according to equation (2.6). The dynamic continues in period 2, with each variable (including y_t) determined in the same manner as in period 1.

2.6. Solution and Simulations

The optimal control model is specified and can now be solved for the optimal path of food aid deliveries. The first step to solving the dynamic optimisation problem is to specify the Hamiltonian.

$$H = \left[(P_t - P_{t-1})^2 + (A - x_t)\right] + \lambda\rho h_t\pi_t\left[\alpha\left(h_t P_t - h_t\pi_t E_{t-1}(P_t)\right) + \beta\left(E_t(P_{t+1}) - P_0\right)\right] \quad (2.15)$$

Recall that the objective function, equation (2.8), is a loss function that we seek to minimise. The first order conditions are the same as in the case of maximising the objective function (a more common practice in optimal control models), but the second order conditions must be investigated. We turn now to the first order conditions, as outlined by the maximum principle.

2.6.1. Solution

The maximum principle states that the optimal solution to the optimal control problem outline in equations above must satisfy the following conditions:

$$\frac{\partial H}{\partial y_t} = 0 \quad (2.16)$$

$$\dot{x}_t = \frac{\partial H}{\partial \lambda_t} \quad (2.17)$$

$$\dot{\lambda}_t = -\frac{\partial H}{\partial x_t}. \tag{2.18}$$

Each of these conditions are addressed in turn. By substituting equations (2.1) and (2.4) into equation (2.15), we produce the reduced-form Hamiltonian:

$$\begin{aligned} H = &\frac{\delta^2 A^2}{(h_t + y_t)^2} - \frac{2\delta A P_{t-1}}{(h_t + y_t)} + P_{t-1}^2 + (A - x_t) + \\ &\lambda\rho\pi_t h_t\left[\alpha\left(h_t \frac{\delta A}{(h_t + y_t)} - h_t\pi_t \frac{\gamma}{h_{t-1}P_{t-1}}\right) + \beta\left(\frac{\gamma(h_t + y_t)}{\delta A h_t} - P_0\right)\right] \end{aligned}. \tag{2.19}$$

The first order partial derivative of the Hamiltonian with respect to the control variable, y_t, is

$$\frac{\partial H}{\partial y_t} = -\frac{2\delta^2 A^2 (h_t + y_t)}{(h_t + y_t)^4} + \frac{2\delta A P_{t-1}}{(h_t + y_t)^2} + \lambda_t\rho h_t\pi_t\left[-\frac{\alpha\delta h_t A}{(h_t + y_t)^2} + \frac{\beta\gamma}{\delta A h_t}\right]. \tag{2.20}$$

Equation (2.20) can be set equal to zero and simplified as

$$\frac{2\delta A P_{t-1} - \alpha\delta\pi_t\rho A\lambda_t h_t^2}{(h_t + y_t)^2} - \frac{2\delta^2 A^2}{(h_t + y_t)^3} - \frac{\beta\gamma\pi\rho\lambda_t}{\delta A} = 0. \tag{2.21}$$

The first order partial derivative of the Hamiltonian with respect to the costate variable, λ, yields the state equation

$$\frac{\partial H}{\partial \lambda_t} = \rho h_t \pi_t \left[\alpha\left(h_t P_t - h_t \pi_t E_{t-1}(P_t)\right) + \beta\left(E_t(P_{t+1}) - P_0\right)\right] = \dot{x}(t)\,. \tag{2.22}$$

Differentiating the negative of the Hamiltonian with respect to the state variable, x, provides the costate equation

$$-\frac{\partial H}{\partial x} = \dot{\lambda}(t) = 1\,. \tag{2.23}$$

The next step in solving for the optimal path is to integrate the costate equation and solve for λ. Using equation (2.23) and the transversality condition in (2.14), we can solve for λ by integrating

$$\lambda = \int_0^T \dot{\lambda}(t)\,dt\,. \tag{2.24}$$

Along the optimal path,

$$\lambda^* = \int \dot{\lambda}\,dt = \int 1dt = t + k\,. \tag{2.25}$$

The constant of integration, k, can be obtained using the transversality condition in (2.14). We know that $\lambda(T) = 0$, so $\lambda(T) = T + k = 0$ and $k = -T$. Therefore, along the optimal path,

$$\lambda_t = t - T. \tag{2.26}$$

Equation (2.26) states that the costate variable is negative throughout the horizon, and is equal to zero at terminal time T. This makes intuitive sense, as planting is beneficial to the aid-recipient country, and therefore the shadow value of planting has a negative impact on the optimised loss function, L.

λ_t can now be substituted into the necessary condition (2.21):

$$\frac{2\delta AP_{t-1} - \alpha\delta\pi_t\rho A(t-T)h_t^2}{(h_t + y_t)^2} - \frac{2\delta^2 A^2}{(h_t + y_t)^3} - \frac{\beta\gamma\pi\rho_t(t-T)}{\delta A} = 0. \tag{2.27}$$

Equation (2.27) satisfies the first order conditions of the dynamic optimal control problem, however we must ensure that the loss function, equation (2.8), is minimised. The second order conditions are established in Appendix A, and they demonstrate that the solution to equation (2.27) minimises the loss function of equation (2.8).

Equation (2.27) can be reduced to a complex cubic polynomial in y_t. An explicit solution for y_t in terms of the parameters and variables (all of which are known values at the time that y_t is decided by the central planner) is troublesome, however. Computation of an explicit solution through Maple mathematics software produces three solutions, all of which contain the square roots of a complicated expression (these solutions can be found in Appendix B). The sign of this complicated expression under the root sign depends on the relative magnitude of the parameters defined in the model. While it is possible to generate positive values for this expression, the parameter values used in the following simulations (see section 2.6.2 for justification of the parameter values used in simulation) make the expression negative and result in polar roots. This situation prevents the simulation of an optimal food aid path as defined by the optimal control model. However simulations can proceed under various other aid paths, according to the behavioural equations of the control model. If an estimated version of this model were to generate parameter values that generated polar root solutions, then an alternative form of the objective function in equation (2.8) could be investigated.[29]

We now turn to comparing the effects of various aid paths on the recipient-country's agricultural industry. The effects of various aid delivery paths on recipient price and production are simulated using the optimal control model that is outlined above.

2.6.2. Simulations

The following simulations utilise generated data. Production and price data from food aid recipient countries is scarce and unreliable. This is especially true in emergency situations, when data collection is made more difficult. Most empirical studies that use

[29] The importance and consequences of the objective function specification are discussed in chapter 2.6.

time series production and price data are modelled in the context of programme and project food aid, for which data are more accessible. Price expectation data is also unavailable. Such data would have to be obtained through a specific survey of a country that receives emergency food aid.

There is, though, only one value that needs to be generated to initiate a simulation of the optimal control system. The level of required nutrition, A, must be stated and all other variables can be produced through the structure and parameters of the model. The economy is presumed to begin in steady-state equilibrium prior to a supply shock, wherein local production equals requirements and there is no emergency food aid. Therefore, $x_{t-1} = h_t = A$ and P_t can be obtained from equation (2.1). Also, $E_t\left(P_{t+1}\right) = P_t$ in the pre-shock state so that $\dot{x} = 0$ prior to the supply shock.

The values of these variables are determined through the model's parameters, which are also generated. As the data is generated through the system's equations, the parameters are not estimated; rather they are assigned values. These assigned values are not arbitrary, however. Some of the parameters can be derived directly from the pre-shock conditions and the others are assigned values that fit logically into the system. A brief description of how each parameter is chosen follows.

ρ

This parameter determines the effect of including $h_t \pi_t$ in the state equation. That is, it affects the speed at which farms can make changes to their output levels. An attempt was made to calibrate this parameter such that a production path that is not degenerate can return so that $x_t \to A$ by the end of a long-run period (which is defined as twenty years[30]).

δ

The price adjustment coefficient determines the food price in the initial steady-state, since the term inside the brackets of equation (2.1) is equal to one before a supply shock. This parameter is set equal to one for clarity of exposition. A larger δ would result in a higher initial price and higher subsequent prices; a smaller δ would produce the opposite response.

γ

The expected price adjustment parameter is difficult to quantify, *a priori*. However, we can use the information provided by the pre-shock condition to calibrate γ such that $E_t\left(P_{t+1}\right) = P_t$ prior to a supply shock. That is, the state equation's expected price effect in the steady state is zero. Equations (2.1) and (2.4) can be set equal to each other and solved for $\gamma = \delta A P_t$.

α and β

These parameters are the most subjective in the system and the relative sizes of the two have the most significant impact on simulated production paths. The size of revenue

[30] Twenty years is chosen to comply with Barrett, Mohapatra and Snyder's definition of a long-term period in food aid deliveries.

and expected price effects determine not only the size of production changes in each period, but also whether production in the recipient country recovers to its pre-shock level or degenerates towards zero. Given the importance of the relative sizes of these two parameters, two separate cases are presented in the following simulations; one in which the revenue effect dominates and one in which the expected price effect dominates.

Though the values of α and β are uncertain, they depend on a few key characteristics of the recipient country. α depends on factors such as the cost of borrowing and on the source of inputs used in planting. As previously mentioned in the discussion regarding the revenue effect, emergency situations may affect a farmer's ability to repay loans from previous periods. The higher is the cost of borrowing, the larger is the amount to be repaid. More funds have to be allocated away from production, hence a larger revenue effect. A farm's source of inputs also affects the size of α. An economy whose farms rely primarily on family labour and do not borrow to acquire capital have a relatively small α. Such farms will not have labour or creditors to repay out of current period's harvest, and a revenue shortfall has a relatively small effect on planting decisions. The opposite is true of farms that rely heavily on hired inputs and repayment of debts out of current period revenues.

A negative supply shock equal to three-quarters of period one's harvest is imposed to initiate the simulation experiment. The result is that $h_1 = \frac{1}{4} x_{t-1} = \frac{1}{4} A$. Supply shocks of different magnitudes were also simulated, and the results are similar to the three-quarters supply shock. The paths of all variables are comparable, but the initial fall in production is larger and the movement to a new steady state is slower. Only the three-quarters supply shock is illustrated below for reasons of brevity.

Within this supply shock scenario, the price and production paths that result from four different aid delivery paths are generated; shortfall aid, excess aid, exogenously-determined aid and exogenously determined aid with a randomised error. A brief discussion of each of these paths is useful.

Shortfall aid

The shortfall aid path delivers aid in the exact amount lost to the exogenous supply shock. The harvest shortfall $(A - h_t)$ is determined and food aid is delivered in that amount. This rule holds in periods following the exogenous supply shock, so that aid continues to be delivered in an amount equal to $(A - h_t)$.

Excess aid

This aid path analyses the effects of the "more is better" philosophy in allocating food aid. The harvest shortfall $(A - h_t)$ in the current period is calculated and aid is delivered in an amount that exceeds that shortfall by ten percent. This path satisfies the "nobody starves" condition, but may have depressing effects on local price and production.

Exogenous aid

This path sets aid deliveries according to a rule unlike either of the above paths. The rule for exogenous aid in the following simulations was arrived upon through a previous permutation of the control model. An earlier version of the model yielded a more tractable formulation for the optimal aid path. This rule, though not necessarily optimal in the context of the current control model, can however be simulated according to the current model's behavioural equations. The rule used in these simulations is

$$y_t = \frac{2\delta A^2}{2\delta A P_{t-1} + T\rho A h_t \pi_t \left(\beta - \alpha h_t\right)} - h_t . \tag{2.28}$$

The rationale for including this aid path in the simulations is to demonstrate that the shortfall aid path is *not* necessarily optimal in either minimising L or in spurring a recovery of local production. This is examined further in the following simulations. Note that in simulations that utilise equation (2.28), the "nobody starves" restriction of equation (2.12) and restriction (2.A.7)[31] are imposed.

Exogenous aid with error

Emergency situations make delivery of food aid logistically difficult. Even if an optimal aid path can be identified, there are several factors that may prevent that optimal amount of aid from reaching its intended recipients. Transportation bottlenecks, hoarding, gangsterism, corruption and weather can all impede aid deliveries. Such complications can result in either too much or too little aid reaching its recipients. The exogenous aid with error simulations include cases where aid is delivered in an amount as determined by equation (2.28), with a random error between

[31] See Appendix A.

plus and minus twenty percent of y_t. Aid is overshot in some periods and undershot in others, depending on the random variable. Note that there is no guarantee that the optimal path with error meets the restriction of equation (2.12); the error may create periods in which the "nobody starves" restriction is violated. Despite this, the price and production paths of this scenario are of interest because of the inherent complications of delivering emergency food aid. The policy implications of delivery errors are discussed in chapter 2.6.3.

All of the above aid paths are analysed under the effects of different relative revenue and expected price effects. One case allows for a relatively large expected price effect and small revenue effect; the other allows for the opposite. Case A provides for a large expected price effect while case B imposes a relatively large revenue effect. Each case is analysed along the four aforementioned aid delivery paths. Each simulation can be categorised by number: scenario A.1 analyses the effects of a shortfall aid path with a relatively large expected price effect. This pattern continues; for example scenario B.3 analyses the effects of exogenously-determined aid with a relatively large revenue effect. The various simulations are summarised in table 2.1.

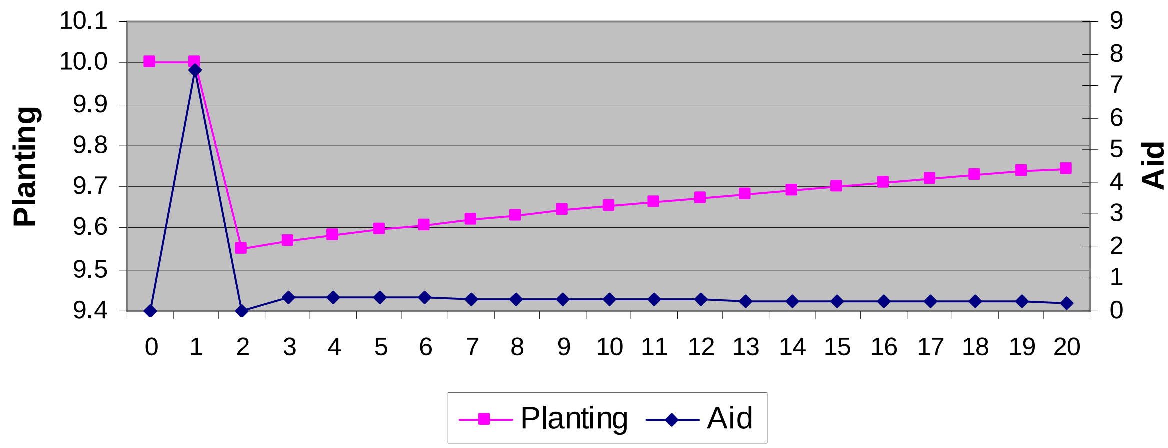
Table 2.1 - Simulation Scenarios

	Shortfall	Excess	Exogenous	Randomised Error
Large expected price effect	A.1	A.2	A.3	A.4
Large revenue effect	B.1	B.2	B.3	B.4

Case A.1 - large expected price effect, shortfall aid

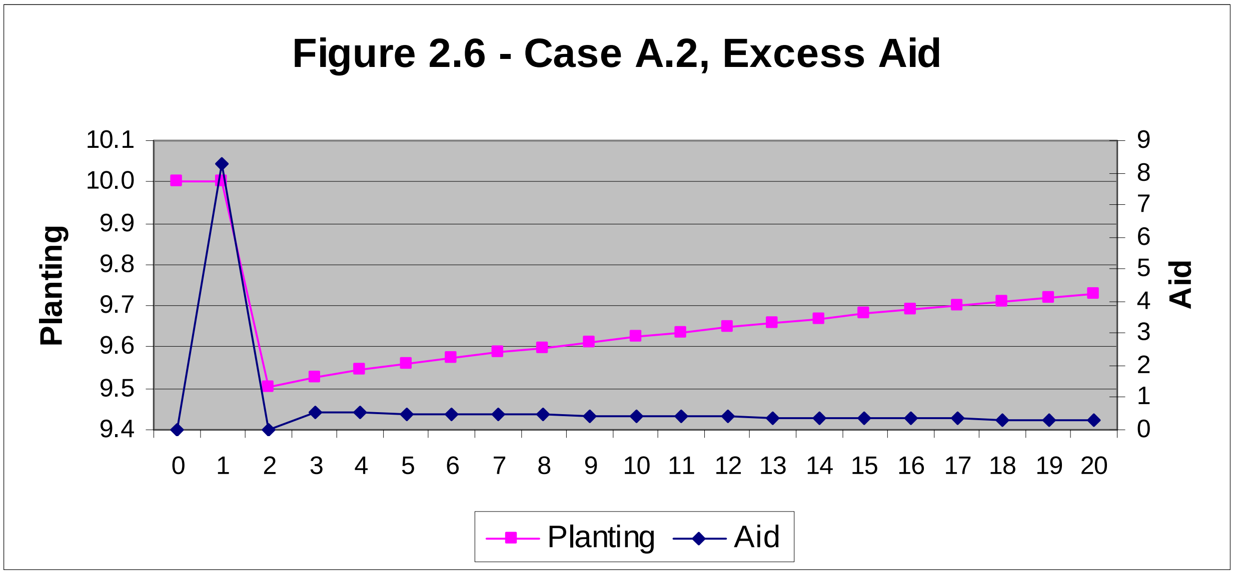
The initial supply shock in period one pushes h_1 down. Aid authorities respond according to the rule $(A - h_t)$ and price adjusts according to equation (2.1). The substantial drop in harvest reduces farm revenue, placing downward pressure on the state equation via the parameter α. However, lower perceived productive capacity results in a rise in price expectation via equation (2.4). Farmers believe that P_2 will rise above P_0 and are motivated to increase production via the expected price effect and parameter β. The net effect on the state equation depends on the relative magnitude of α and β, which in this case favours β. The state equation assumes a positive sign, however x_1 remains at A due to the constraint in equation (2.10). This is not the end of the story, however. Recall that the farmer expects price in period two to rise; it is this expectation that pushes $\dot{x}$ above zero in period one. The higher expected price is how the farmer expects to recoup the costs of having increased production in period one above what period one's revenue would dictate. P_2, however, stays below $E_1(P_2)$ and produces a negative revenue effect in period two; the result is decreased planting in period two. From this point on, the revenue effect remains negative, but is dominated by the expected price effect. Production recovers towards its pre-shock level, A, and food aid shipments fall towards zero by the end of the 20 period simulation. The cumulative loss function is 6.49. Case A.1 is illustrated in figure 2.5.

Figure 2.5 - Case A.1, Shortfall Aid



Case A.2 - large expected price effect, excess aid

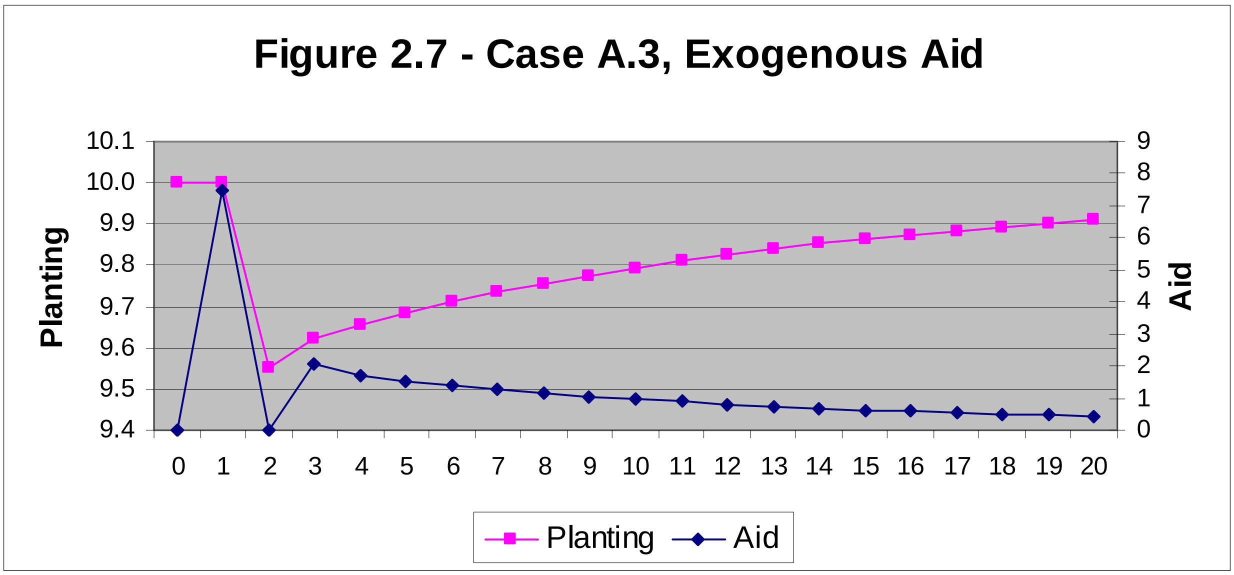
This scenario delivers food aid in an amount equal to the crop shortfall ($A - h_t$) plus an additional ten percent in each period. The production path in this scenario is similar to the shortfall aid path, but differs in two key ways. First, the negative revenue effect is larger in period two, resulting in lower planting in period two. The second difference is the speed at which local production recovers. x_t is further below A at the end of the twenty-period simulation in the excess aid scenario than in the shortfall aid scenario. The cumulative loss function is 7.04. Case A.2 is illustrated in figure 2.6.



Case A.3 - large expected price effect, exogenous aid

The motivation for including a simulation of this aid path (as outlined in equation (2.28)) is to demonstrate that the shortfall aid path is not necessarily optimal. The planting and aid trajectories look similar to the shortfall aid case, and are identical in certain periods because the "nobody starves" constraint binds. Local production recovers more quickly, however, in the exogenous aid case than in the shortfall aid case. It is interesting to note that there are several periods in which this aid path is above the shortfall path, indicating that just meeting the "nobody starves" condition is not necessarily optimal.

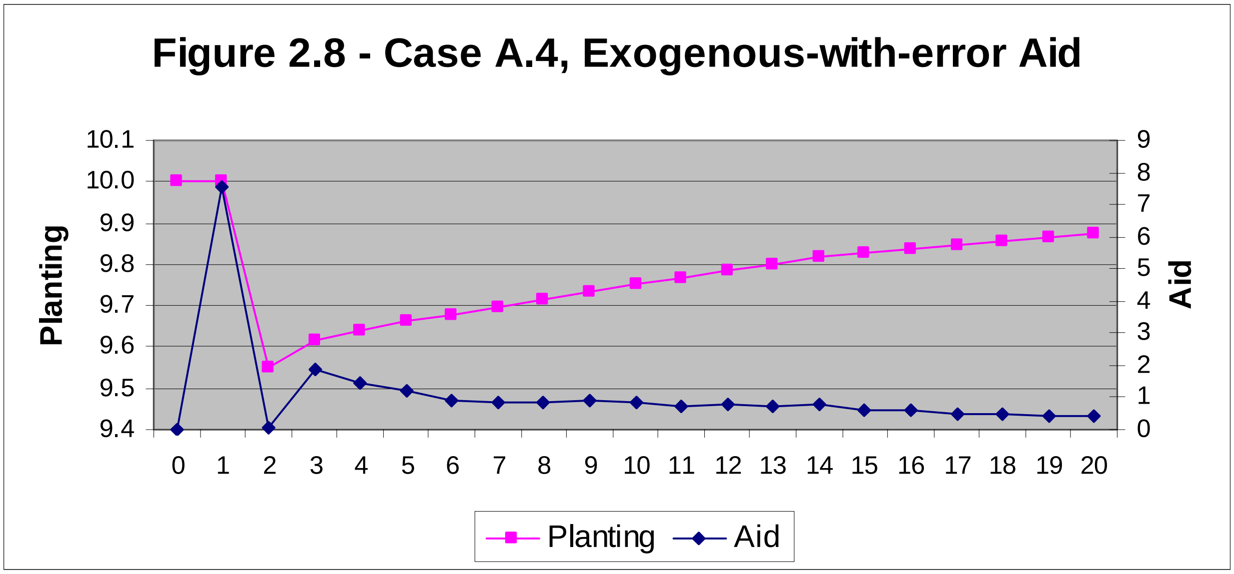
It seems logical, *a priori*, that replacing just the amount of food lost to the shock should minimise damage to the recipient economy. In the case of a large expected price effect, the exogenous amount of aid is equal to the lost harvest in the first two periods, and is above the lost harvest for the next eight periods. It seems as though a higher level of food aid keeps price below what would be the case in the $(A - h_t)$ path. This results in a larger expected price effect, and faster production recovery. The cumulative loss function in the case of exogenous aid is 4.09. Case A.3 is illustrated in figure 2.7.



Case A.4 - large expected price effect, exogenous aid with a randomised error

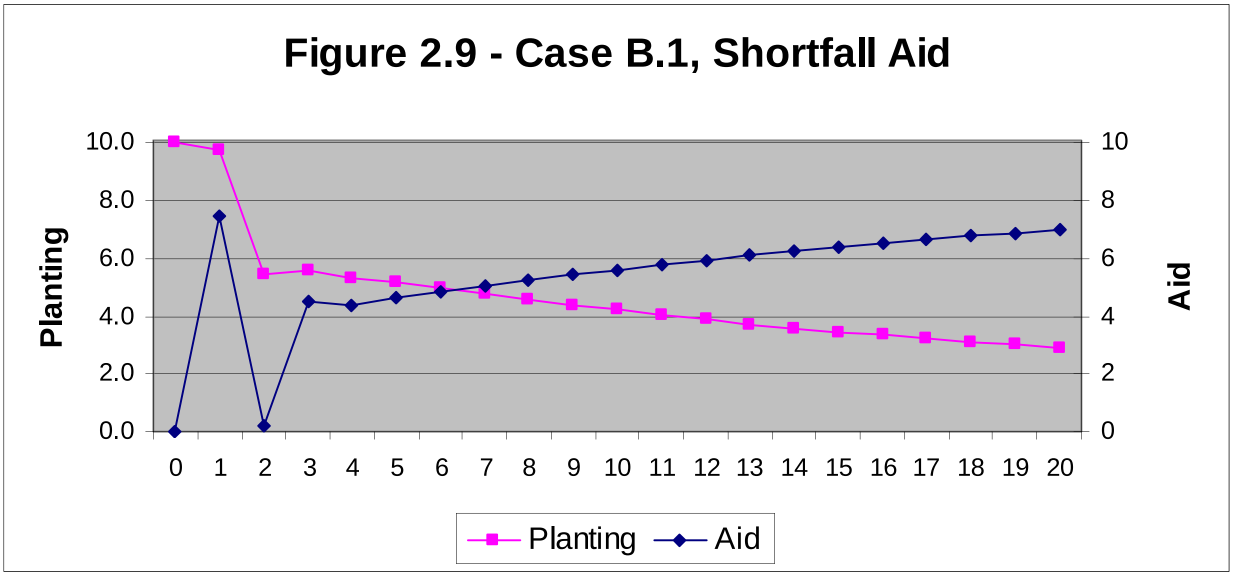
The production path that results from exogenous aid with a randomised error closely resembles case A.3. The initial revenue effect produces a larger decrease in period two planting, however, since a positive error in aid delivery reduces current price. Also, local production recovers more slowly when aid is delivered with errors. x_t is nearer A after twenty periods in scenario A.3 than in scenario A.4. The cumulative loss function is 4.70. Case A.4 is illustrated in figure 2.8.

Note that there are an unlimited number of potential paths that scenario A.4 could take, depending on the direction and magnitude of the random error. Such paths could (and likely would) include negative delivery errors, thus violating the "nobody starves" condition.



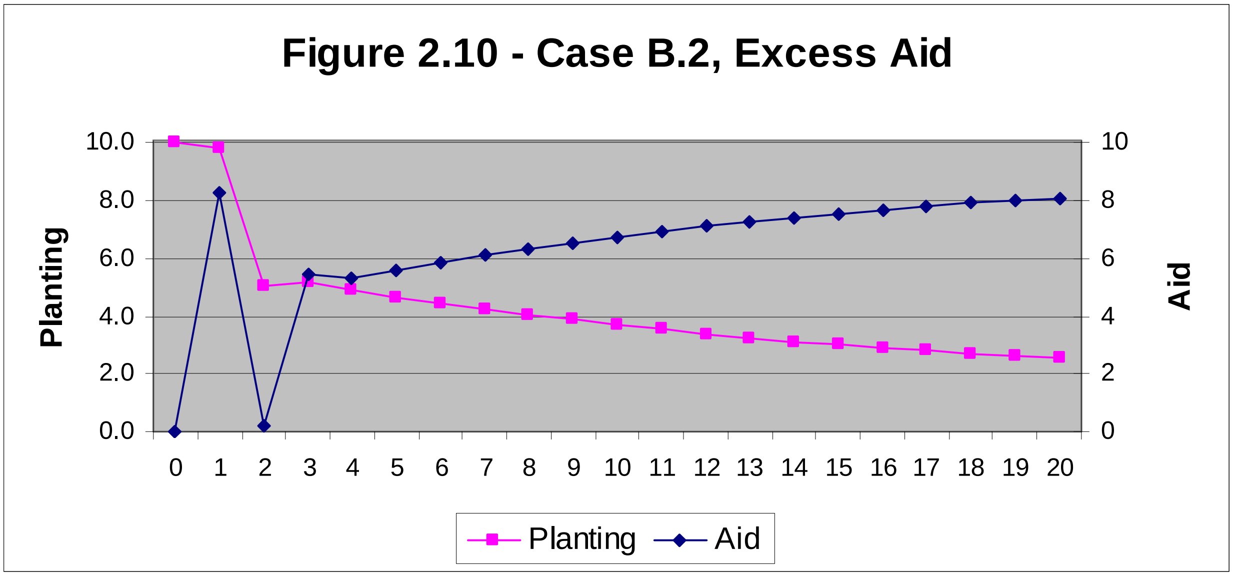
Case B.1 - large revenue effect, shortfall aid

A relatively large revenue effect results in a markedly different production path in the case of a negative supply shock. The initial supply shock in period one reduces current revenue, which dominates the positive expected price effect and pushes current planting down. However unlike Case A.1, production does not recover in period three. The negative revenue effect continues to dominate the positive price effects and the state equation is persistently negative. Local planting degenerates towards zero, and aid shipments rise towards A. The ultimate steady-state result (beyond the twenty-period simulation) would be complete dependence on food aid. The cumulative loss function is 111.41. Case B.1 is illustrated in figure 2.9.



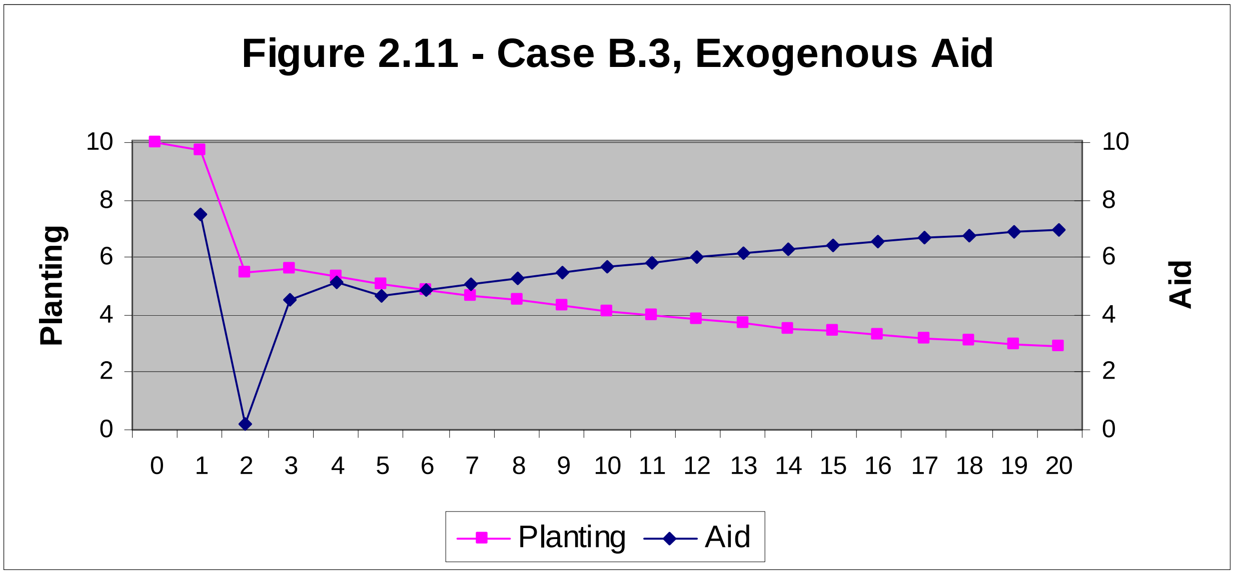
Case B.2 - large revenue effect, excess aid

The trends in this simulation are similar to case B.1, but local planting degenerates at a faster rate. The initial negative revenue effect is larger and the continued dominance of the negative revenue effect results in relatively sharp decreases in domestic planting and correspondingly fast increases in food aid deliveries. The cumulative loss function is 120.09. Case B.2 is illustrated in figure 2.10.



Case B.3 - large revenue effect, exogenous aid

The production trajectory generated by the aid path outlined in equation (2.28) is similar to that of case B.1, however unlike the scenarios that include a large expected price effect, the exogenous aid path is not superior (in terms of generating a recovery path for local planting or minimising the objective function) to the shortfall aid path. This indicates that the optimal path (as defined by the solution to equation (2.27)) depends on the relative magnitudes of α and β. The cumulative loss function is 112.12. Case B.3 is illustrated in figure 2.11.



Case B.4 - large revenue effect, exogenous aid with a randomised error

The randomised delivery error creates a faster degenerative production path than does the exogenous aid path. Domestic planting falls towards zero more rapidly than in case B.3; local planting is nearer zero at the end of the twenty-period simulation in case B.4 than in case B.3. Recall that there exist an unlimited number of potential paths that scenario B.4 could take, depending on the size and direction of the random error. Also, some periods experience a negative delivery error, resulting in a violation of the "nobody starves" condition. The cumulative loss function is 126.81. Case B.4 is illustrated in figure 2.12.

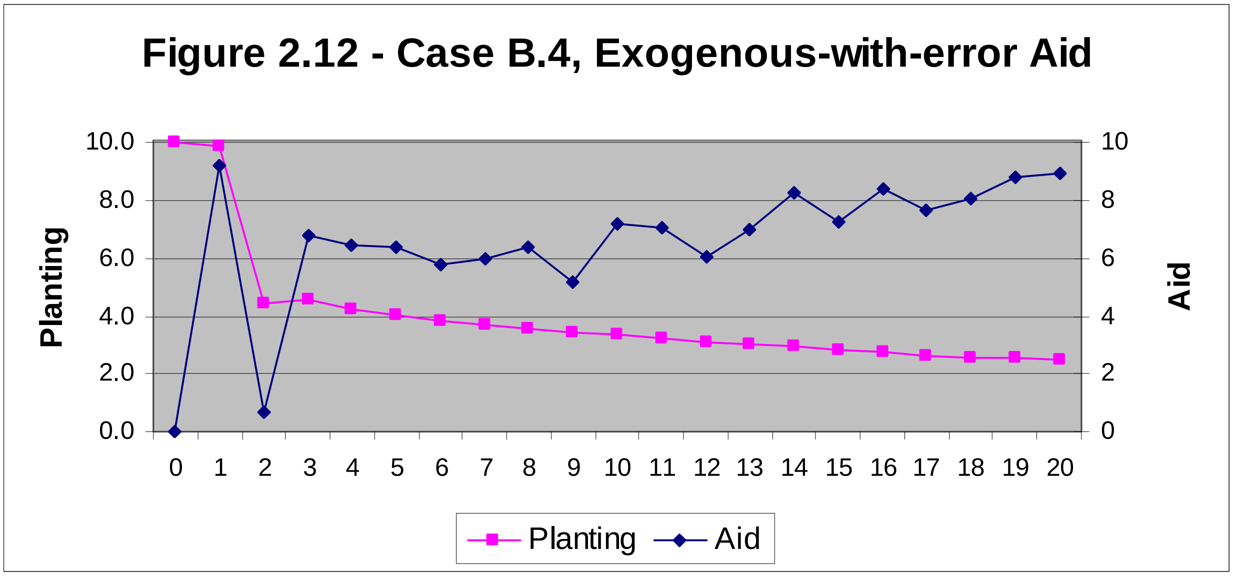


Table 2.2 summarises the simulation results. All aid paths result in a production recovery in cases of a relatively large expected price effect. The exogenous aid path generates the smallest cumulative loss function and the excess aid path generates the largest loss function. In cases of a relatively large revenue effect, none of the simulated aid paths allow local production to recover. Shortfall aid generates the smallest cumulative loss function and exogenous aid with error generates the largest cumulative loss function.

Table 2.2 - Simulation Results

	Shortfall	Excess	Exogenous	Randomised Error
Large expected price effect $\alpha = 0.01, \beta = 0.3$	Case A.1 $L = 6.49$ recovery	Case A.2 $L = 7.04$ recovery	Case A.3 $L = 4.09$ recovery	Case A.4 $L = 4.70$ recovery
Large revenue effect $\alpha = 0.1, \beta = 0.2$	Case B.1 $L = 111.41$ degeneration	Case B.2 $L = 120.09$ degeneration	Case B.3 $L = 112.12$ degeneration	Case B.4 $L = 126.81$ degeneration

2.6.3. Discussion

The simulations in this essay are based on generated data and parameters and, as such, cannot be directly applied to a specific case of emergency food aid. There are, however, several important insights that can be gleaned from the model and its simulation results.

First, accurate measures of the model's relevant variables and constants must be obtained to successfully manage a food emergency. Planners should know the size of an affected population, and how much food is required to sustain their health through the course of the emergency; that is, there should be an accurate estimate of A. The scale of the supply shock must also be considered in any food aid response. A comparison of harvest sizes before and after a supply shock (h_t) would reveal this information. Accurate price data are also necessary to measure the effects of aid on local markets.

Another insight provided by this model is the importance of the initial supply shock's impact on revenue and resulting production decisions. This issue has been largely ignored in food aid analysis, but is key in determining the path that production takes in response to an emergency situation and subsequent aid deliveries. The revenue effect demonstrates that even if there exists sufficient price incentives to increase production, farmers may be unable to do so because of insufficient revenue.

The importance of the revenue effect must be compared to the effects of expected price effects. Planners must have an understanding of how farmers' price expectations are formed, and how those expectations affect planting decisions. As the previous simulations demonstrate, the relative sizes of these two effects can determine whether local production recovers or degenerates.

If planners can obtain reliable estimates of the data and parameters described above, then some policy conclusions can be deduced. If the expected price effect dominates the revenue effect, then the simulations show that local production is likely to recover after an emergency. If an optimal aid path can be calculated, then this path will minimise damage to the recipient-country's agricultural industry. There may, however, be an important reason to stray from this optimal path. The realities of delivery disruptions (as illustrated in cases A.4 and B.4) create the possibility of aid shortages in certain regions. A prudent response would be to add a fixed amount of aid to the optimal delivery path in each period to compensate for negative delivery errors. Doing so may slow the recovery, but save lives in the process.

A larger policy problem emerges if the revenue effect dominates the expected price effect. The simulations show that all constructed aid paths result in degenerative production paths[32]. The problem is that the negative impact of the revenue effect on farmers' planting decisions perpetually dominates any positive impact that the expected price effect might have. Recall that the importance of the revenue effect (i.e. the size of

[32] A scenario that included zero aid in every period was simulated in the context of a relatively large revenue effect. Production recovered towards A in this case, but the simulation was not included because of its obvious violation of the "nobody starves" requirement.

α) will depend on the characteristics of the recipient country's agricultural industry. For example, an agricultural region that relies relatively heavily on hired labour and inputs is likely to have a larger α than a region that uses primarily family labour.

One possible policy response in the case of a large revenue effect would be to directly subsidise farmers in emergency situations. This could help to increase farmers' current revenue to be either equal to or above expected revenue. The positive impact of expected price effects can then be realised and even complimented. Another option is to provide some aid in the form of variable inputs, such seeds or fertilisers. Farmers that could have otherwise not afforded such inputs (due to revenue shortfalls) can use these inputs to increase current-period planting.

It is important to remember that the model presented in this paper does not incorporate valuation of lives lost to starvation. This is because such loss is assumed to trump any possible negative market effects that aid might have on a recipient country. Emergency food aid should always be delivered in an amount sufficient to prevent starvation. Lower levels of aid delivery may result in faster recovery for the agricultural sector and a lower value of the model's value function, but such paths do not meet humanitarian need requirements. The primary goal of any emergency food aid plan should be to avert starvation, the secondary goal to minimise potential damage to the recipient-country's agricultural industry.

As in any optimal control model, the conclusions and results of the solution and simulations are sensitive to the specification of the objective function. The loss function of equation (2.8) does not weight the relative importance of price volatility versus the importance of deviations from capacity output. Both terms enter into the loss function with no attached parameter. Weighting one of the components of the objective function differently would produce a different optimal path. Because the control model is simulated, and not estimated, a decision was made to leave the objective function unweighted for generality. An empirical application of this model would warrant investigation of various forms of the objective function in equation (2.8) in order to generate alternative optimal aid paths.

Also, the importance of the "nobody starves" condition cannot be overstated in the simulation results. Violation of the "nobody starves" condition would cause a food shortage, and result in higher food prices. This will turn revenue effects positive, placing upward pressure on current period planting. However, this model is built on the principle that food aid should meet the needs of the recipient country. Allowing for an acute food shortage does no such thing. Also, an acute food shortage in which people starve will have negative effects on the affected region's productive capacity. Labourers that do not receive enough food will either die or become less productive. The current model has no mechanism to account for reduced productive capacity as the result of dying farmers. Note, however, that while violating the "nobody starves" condition may generate a faster recovery path for local planting, such a violation will not necessarily result in a lower cumulative loss function. The price volatility that results from an acute food shortage can push the objective function above that of the shortfall aid path value. Again, this result is dependent on the specification and weighting of the objective function.

Beyond the attempt to best meet the needs of the recipient country, this research provides an avenue for further consideration in the formulation of trade rules regarding food aid. Current WTO trade negotiations could benefit from the identification of an amount of aid that a recipient country needs. Surplus disposal of agricultural commodities is a contentious issue in WTO negotiations, and the current rules determine the legitimacy of food aid from the donor-country's perspective. Shifting the determination of aid's legitimacy to the recipient country's perspective would provide the WTO with more sensible guidelines by which to govern food aid. If a needed, or optimal, amount of aid can be determined then the WTO should allow for aid shipments up to (and perhaps just beyond) that amount. Any aid above that amount should raise concern for two reasons. The first reason is that less aid will be able to meet the region's nutritional requirements without imposing as large a hardship on local agriculture. A second reason is that food aid shipments that exceed the needed quantity may be an attempt by donor countries to satisfy domestic concerns, specifically surplus disposal. The WTO should be particularly interested in the second reason.

2.7. Conclusions

The primary goal of food aid deliveries should be to benefit the recipient country. The experience over the fifty-odd years of government-funded food aid programmes has, however, not always been so philanthropic. US food aid programmes were initiated primarily as a means of surplus disposal, and this tradition has continued for more than fifty years. Programme and project food aid have been shown to be linked more closely with donor motives than with recipient needs.

The possibility that food aid shipments may have negative effects on recipient country agricultural industries, first analysed by Schultz, has generated a sizable stream of literature. This literature has tried to qualify and quantify the effects that food aid might have on a recipient country. A surprising omission from this literature is an attempt to identify an amount of food aid that would provide the most benefit to the recipient country. The current research tries to make up for this omission by defining how much emergency food aid a recipient country needs, and building a control model to identify that need.

Identifying the amount of food aid that most benefits a recipient country can provide the WTO with a framework by which to evaluate the legitimacy of food aid. Food aid shipments that arrive in amounts large enough to depress local production beyond a minimum amount should be viewed with caution.

This research provides a new perspective on food aid by formulating a definition of need in the context of emergency food aid and constructing an optimal control model around that definition. The control model is solved for the optimal quantity of food aid shipments. Though the control solution is not simulated, several important conclusions can be drawn from the control model and its simulations.

First, food aid planners must have accurate estimates of the model's important variables and parameters. Specifically, planners must know an affected region's nutritional requirements and the size of the nutritional shortfall that is caused by a supply shock. Planners must also have an understanding of how farmers react to changes in current revenue and expected price effects. The relative magnitude of these effects can determine whether local production recovers or degenerates.

Second, any food aid delivery plan must consider the possibility of delivery obstacles. Food aid shipments, especially in emergency situations, are unlikely to reach recipients in the exact amount intended by donors. One way to deal with such obstacles is to add a buffer stock to aid shipments to account for possible negative delivery errors. A slower production recovery might be a reasonable price to pay for saved lives.

Third, the model's simulations show that *less* is not always *more* in the case of emergency food aid. Specifically, the simulations demonstrate that an amount of aid that just makes up for a region's nutritional shortfall will not necessarily result in the fastest recovery path for local production. An exogenous aid path was shown to result in faster local production and to produce a smaller value for the objective loss function.

There are several avenues that further research could take. One possible step is to include an expectation of future food aid in farmers' price expectation formulation. There exists some evidence that programme food aid is persistent (Barrett, Mohapatra and Snyder). Equation (2.4) could look something like

$$E_t\left(P_{t+1}\right) = \frac{\gamma}{h_t P_t} + f\left(y_t\right). \tag{2.29}$$

Allowing farmers to expect food aid in the next period would have two primary effects. First, it will alter their expectation of future price, thereby affecting the expected price effect in the current planting decision. Also, farmers' expectation of aid will affect the size of the revenue effect; if more aid was delivered than was expected, then current price will be lower than expected price and a revenue shortfall will result.

Another path that this research could take would be to allow for a measure of the number of people at risk of starvation during various sizes of supply shocks. The model could then simulate the marginal effects of increasing food aid shipments on the number of at risk people. A benefit-cost analysis could then be presented that weighs the costs of increasing emergency aid against the benefits of saving an additional life. Such an analysis would make for an interesting policy discussion.

APPENDIX A - SUFFICIENCY CONDITIONS

There exist several methods for establishing the sufficiency of the maximum principle, depending on the structure of the optimal control problem. The method outlined on page 163 of Leonard and Van Long is used below.

To establish that the path suggested by equation (2.27) minimises the cumulative loss in equation (2.8), the following proposition must be established:

> If L of equation (2.8) is convex in (x, y) jointly, and $\lambda \leq 0$ and $\dot{x}$ of equation (2.6) is convex in (x, y) jointly, then the necessary conditions of the simplified maximum principle are sufficient for an optimal solution in minimising L.

1. Convexity of L

$$\frac{\partial L}{\partial x} = -1 \tag{2.A.1}$$

$$\frac{\partial^2 L}{\partial x^2} = 0 \tag{2.A.2}$$

$$\frac{\partial^2 L}{\partial x \partial y} = \frac{\partial^2 L}{\partial y \partial x} = 0 \tag{2.A.3}$$

$$\frac{\partial L}{\partial y} = \frac{2\delta A P_{t-1}}{\left(h_t - y_y\right)^2} - \frac{2\delta^2 A^2}{\left(h_t + y_t\right)^3} \tag{2.A.4}$$

$$\frac{\partial^2 L}{\partial y^2} = \frac{-4\delta A P_{t-1}\left(h_t + y_t\right)}{\left(h_t + y_t\right)^4} + \frac{6\delta^2 A^2 \left(h_t + y_t\right)^2}{\left(h_t + y_t\right)^6} \tag{2.A.5}$$

The Hessian matrix of L is

$$Hessian = \begin{bmatrix} \dfrac{2\delta A\left(3\delta A - 2P_{t-1}\left(h_t + y_t\right)\right)}{\left(h_t + y_t\right)^4} & 0 \\ 0 & 0 \end{bmatrix}. \tag{2.A.6}$$

The Hessian is positive semidefinite if the term $\dfrac{2\delta A\left(3\delta A - 2P_{t-1}\left(h_t + y_t\right)\right)}{\left(h_t + y_t\right)^4}$ is greater than or equal to zero. This inequality can be solved for y_t to provide a restriction on aid deliveries.

$$y_t \le \left(\frac{3}{2}\frac{\delta}{P_{t-1}}\right)A - h_t \tag{2.A.7}$$

Inequality (2.A.7) states that the value function, L, is convex in x and y when this inequality is satisfied.

2. Convexity of $\dot{x}$

$$\frac{\partial \dot{x}}{\partial x} = 0 \tag{2.A.8}$$

$$\frac{\partial^2 \dot{x}}{\partial x^2} = 0 \tag{2.A.9}$$

$$\frac{\partial^2 \dot{x}}{\partial x \partial y} = \frac{\partial^2 \dot{x}}{\partial y \partial x} = 0 \tag{2.A.10}$$

$$\frac{\partial \dot{x}}{\partial y} = \rho\pi h_t\left[\frac{-h_t\alpha\delta A}{(h_t + y_t)^2} + \frac{\beta\gamma}{\delta A h_t}\right] \tag{2.A.11}$$

$$\frac{\partial^2 \dot{x}}{\partial y^2} = \rho\pi h_t\left[\frac{2\alpha\delta A h_t (h_t + y_t)}{(h_t + y_t)^4}\right] \tag{2.A.12}$$

The Hessian matrix of $\dot{x}$ is

$$Hessian = \begin{bmatrix} 0 & 0 \\ 0 & \rho\pi h_t\left[\frac{2\alpha\delta A h_t (h_t + y_t)}{(h_t + y_t)^4}\right] \end{bmatrix}. \tag{2.A.13}$$

The term $\rho\pi h_t\left[\frac{2\alpha\delta A h_t (h_t + y_t)}{(h_t + y_t)^4}\right]$ is positive, so the Hessian matrix of $\dot{x}$ is positive semidefinite and $\dot{x}$ is convex in (x, y).

3. Sign of λ_t

Equation (2.26) shows that $\lambda_t = t - T$, so that λ_t is less than zero in all periods other than T, when $\lambda_t = 0$.

Therefore, L of equation (2.8) is convex in (x, y) jointly, and $\lambda \leq 0$ and $\dot{x}$ of equation (2.6) is convex in (x, y) jointly. The necessary conditions of the simplified maximum principle are sufficient for an optimal solution in minimising L.

APPENDIX B - SOLUTIONS TO THE OPTIMAL CONTROL PROBLEM

Three solutions to equation (2.27) are provided by computation in Maple. They are:

$$y_t^1 = \frac{1}{\beta\delta\rho\lambda_t}\left\{\delta A\left[\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_tP_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right\}$$

$$-\frac{\left(\delta A\left(\alpha\rho\lambda_t h_{t-1}^2 + 2\pi_tP_{t-1}\right)\right)}{\left\{\left[\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_tP_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right\}} - h_t$$

(2.A.14)

$$y_t^2 = -\frac{1}{6\beta\delta\rho\lambda_t}\left\{\delta A\left[\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_tP_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right\} +$$

$$\frac{\left(\delta A\left(\alpha\rho\lambda_t h_{t-1}^2 + 2\pi_tP_{t-1}\right)\right)}{\left\{\left[2\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_tP_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right\}} - h_t +$$

$$\frac{1}{2}\pi_t\sqrt{3}\left\{\frac{1}{\beta\gamma\rho\lambda_t}\left[\delta A\left[\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_tP_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right]\right\}$$

$$+\frac{\left(\delta A\left(\alpha\rho\lambda_t h_{t-1}^2 + 2\pi_tP_{t-1}\right)\right)}{\left\{\left[\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_tP_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right\}}$$

(2.A.15)

$$
\begin{aligned}
y_t^3 = &-\frac{1}{6\beta\delta\rho\lambda_t}\left\{\delta A\left[\frac{\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_t P_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)}{\beta^2\gamma^2\rho^2\lambda_t^2}\right]^{\left(\frac{1}{3}\right)}\right\} + \\
&\frac{\left(\delta A\left(\alpha\rho\lambda_t h_{t-1}^2 + 2\pi_t P_{t-1}\right)\right)}{\left\{\left[2\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_t P_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right\}} - h_t - \\
&\frac{1}{2}\pi_t\sqrt{3}\left\{\frac{1}{\beta\gamma\rho\lambda_t}\left[\delta A\left[\frac{\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_t P_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)}{\beta^2\gamma^2\rho^2\lambda_t^2}\right]^{\left(\frac{1}{3}\right)}\right]\right\} \\
&+ 3\sqrt{3}\,\frac{\left(\delta A\left(\alpha\rho\lambda_t h_{t-1}^2 + 2\pi_t P_{t-1}\right)\right)}{\left\{\left[\left(27\pi_i + 3\sqrt{3}\sqrt{\frac{\alpha^3\rho^3\lambda_t^3h_t^3 + 6\pi_t\alpha^2\rho^2\lambda_t^2h_t^4P_{t-1} - 12\alpha\rho\lambda_t h_t^2P_{t-1}^2 - 8\pi_t P_{t-1}^3 - 27\beta\gamma\rho\lambda_t}{\beta\gamma\rho\lambda_t}}\right)\beta^2\gamma^2\rho^2\lambda_t^2\right]^{\left(\frac{1}{3}\right)}\right\}}
\end{aligned}
$$

(2.A.16)

REFERENCES

Barrett, C. B. "Food Aid: Is It Development Assistance, Trade Promotion, Both, or Neither." *American Journal of Agricultural Economics* 80(1998): 566-571.

Barrett, C. B., S. Mohapatra, and D. L. Snyder. "The Dynamic Effects of US Food Aid." *Economic Inquiry* 37, no. 4(1999): 647-656.

Bezuneh, M., B. Deaton, and S. Zuhair. "Food Aid Disincentives: the Tunisian Experience." *Review of Development Economics* 7, no. 4(2003): 609-621.

Diven, P. J. "The domestic determinants of US food aid policy." *Food Policy* 26(2001): 455-474.

Donovan, Cynthia, Robert J. Myers, David L. Tschirley and Michael T. Weber. (1999) The Effects of Food Aid on Maize Prices in Mozambique. Sacramento, CA, Aldershot, Hants, UK Ashgate.

Fisher, F. M. "A Theoretical Analysis of the Impact of Food Surplus Disposal on Agricultural Production in Recipient Countries." *Journal of Farm Economics* 45, no. 4(1963): 863-875.

Granger, C., and P. Newbold. "Spurious Regressions in Econometrics." *Journal of Econometrics* 2(1974): 111-120.

Gray, Richard, "Economic Measures of Sustainability." *Canadian Journal of Agricultural Economics* 39, no. 4(1991): 627-635.

Isenman, P. J., and H. W. Singer. 2001. "Food Aid: Disincentive Effects and Their Policy Implications", in V. Ruttan (ed.), Why Food Aid? The Johns Hopkins University Press, Baltimore.

Kamien, M. I., and N. L. Schwartz. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. 2nd Edition ed. Vol. 31. Advanced Textbooks in Economics. Edited by C. J. Bliss, and M. D. Intriligator. New York: North-Holland, 1991.

Kandiero, T., and M. Chitiga (2003) Trade Openness Foreign Direct Investment in Africa. Cape Town, South Africa.

Kenyon, D. E. "Producer Ability to Forecast Harvest Corn and Soybean Prices." *Review of Agricultural Economics* 23, no. 1(2001): 151-162.

Leonard, D., and N. Van Long. *Optimal control theory and static optimization in economics*. Cambridge: Cambridge University Press, 1992.

Lowder, S. K. "A Post-Schultzian View of Food Aid, Trade, and Developing Country Ceral Production: A Panel Data Analysis." PhD, The Ohio State University, 2004.

Mann, J. S. "The Impact of Public Law 480 on Prices and Domestic Supply of Cereals in India." *American Journal of Agricultural Economics* 49(1968): 131-146.

Maxwell, S. J., and H. W. Singer. "Food Aid to Developing Countries: A Survey." *World Development* 7(1979): 225-247.

Rogers, K. D., U. K. Srivastava, and E. O. Heady. "Modified Price, Production and Income Impacts of Food Aid Under Market Differentiated Distribution." *American Journal of Agricultural Economics* 54(1972): 201-208.

Rothschild, E. 1993. "Is it Time to End Food For Peace?", in V. Ruttan (ed.), Why Food Aid? The Johns Hopkins University Press, Baltimore.

Ruttan, V. W. "Does Food Aid Have a Future?" *American Journal of Agricultural Economics* 80, no. 3(1998): 572-574.

Schultz, T. W. "Value of US Farm Surpluses to Underdeveloped Countries." *Journal of Farm Economics* 42, no. 4(1960): 1019-1030.

Sen, A. *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford: Clarendon Press, 1981.

Shaw, D. J., and H. W. Singer. "A future food aid regime: implications of the Final Act of the Uruguay Round." *Food Policy* 21, no. 4/5(1996): 447-460.

Shideed, K. H., and F. C. White. "Alternative Forms of Price Expectations in Supply Analysis For US Corn and Soybean Acreages." *Western Journal of Agricultural Economics* 14, no. 2(1989): 281-292.

Singer, H., J. Wood, and T. Jennings. *Food Aid: The Challenge and the Opportunity*. Oxford: Clarendon Press, 1987.

Srinivasan, T. N. "Food Aid: A Cause of Development Failure or and Instrument of Success?" *The World Bank Economic Review* 3, no. 1(1989): 39-65.

Varian, H. R. *Microeconomic Analysis*. 3rd Edition ed. New York: W. W. Norton & Company, 1992.

Zahariadis, N., R. Travis, and J. B. Ward. "US food aid to Sub-Saharan Africa: politics or philanthropy?" *Social Science Quarterly* 81, no. 2(2000): 663-676.

ESSAY 3: THE IMPORTANCE OF DYNAMICS IN COMMODITY ADVERTISING DEMAND SYSTEMS

3.1. Introduction

Measuring the effectiveness of advertising is one of the most active research areas in agricultural economics. The large sums of money, both private and public, spent on agricultural promotion are testaments to the importance of accurately assessing its effectiveness. Hayes estimates that for every US$500 million spent each year on US agricultural promotion programs, the average cost to farmers is US$1000. Canada has similar compulsory producer-level programs. Such programs are often mandatory, in that producers of a promoted good must contribute to the marketing programs through a "check-off" system. The mandatory nature of such programs combined with the fact that public funds are also allocated to agricultural promotion programs, heightens the public interest in accurate appraisals of promotion programs.

The majority of attempts to assess advertising effectiveness do so by estimating demand equations or systems and testing the impact of advertising expenditures. Results from such analyses are inconsistent and often counterintuitive. Different functional forms yield different conclusions using similar data, and elasticity estimates are frequently counterintuitive. Furthermore, static demand-based studies overlook dynamic market interactions that can influence the effectiveness of advertising over time. Short-run elasticities are sure to be different from long-run elasticities, and basing policy decisions on short-run elasticities may be short sighted.

Another potentially serious drawback of current advertising assessment methodologies is the neglect of the underlying data's time-series properties. Time-series data used in demand model estimation is likely to be nonstationary[33]; ignoring this information in estimation may lead to spurious regression and misleading model inference.

A third shortcoming of current methodologies is that, if data are nonstationary, then modelling demand in anything other than error-correction form may omit some important information about long-run relationships between the model's constituent variables. This information is important for two reasons. First, if variables are cointegrated then estimating a model without an error-correction term amounts to a misspecification error (Enders). Second, it seems likely that the difference between short and long-run elasticities will be more significant in the case of products with close substitutes. The prices of such closely substitutable products may share a long-run

[33] See chapter 3.4.2 for an explanation of stationarity,.

relationship, so investigating cointegration within the context of a demand model for such products is important.

This essay explores the consequences of failing to account for potential long-run and dynamic effects in a demand system and explains the importance of correctly modelling the time-series properties of the data used in demand models. This explanation includes reasons that consideration of time-series properties is particularly important in advertising-demand models of closely substitutable products. A demand system that correctly tests for and accommodates dynamic and time-series properties is developed and applied to US meat data. The results of this model are compared to a traditional, static demand system.

This introductory chapter is followed by seven chapters. Chapter 3.2 provides background on commodity promotion and its importance in agriculture. Chapter 3.3 reviews the methods by which the effectiveness of advertising has been assessed in economic studies and includes comments on the shortcomings of current methodologies. Chapter 3.4 explains the importance of considering the dynamic and time-series properties of the data used in demand models and explains why such properties might be particularly relevant in modelling demand for products that have close substitutes. Chapter 3.5 builds an empirical model that accounts for dynamic and time-series properties, and chapter 3.6 estimates this model using US meat data. A traditional, static model is also estimated and the results of the two models are compared. Chapter 3.7 discusses the use of demand models and their elasticities in policy decisions and chapter 3.8 closes with some concluding remarks and suggestions for further research.

3.2. Commodity Promotion

3.2.1. The Rationale for Advertising

Advertising expenditures are made in hopes of generating higher profits. A shift in consumer preferences is the mechanism by which advertisers hope to increase profits. The rationale is as follows: advertisers undertake campaigns to convey “new” information to consumers in hopes of altering their preferences. New preferences are reflected by a new (shifted out) demand curve for the advertised product. The resulting higher price translates into higher profits for producers. A simplified static representation of this effect is illustrated in figure 3.1. Consider a market with a representative firm whose advertising expenditures successfully alter consumer tastes so that short-run demand shifts from D_{SR}^{0} to D_{SR}^{1}. The new higher price provides higher profits to producers. This representation is simplified by not considering the costs of the advertising campaign or the (perhaps rising) costs of increasing output.

There exist two broad categories of advertising, reflecting the nature of the products; branded advertising and generic advertising. Branded advertising generally refers to advertising programs in markets that feature non-homogenous products. Products such as cars and clothing may be sufficiently differentiable so that an advertising campaign for one brand does not necessarily increase demand for all brands across the group of products. Most branded advertising is done at the firm level since individual firms have private incentives (higher profits) to shift out the demand curve for their products.

Generic advertising describes promotional activity in markets that are comprised of homogenous, highly substitutable products. Often, such products are not branded, so consumers are not inclined to choose one variety over another. The benefits (measured in terms of higher prices) of generic advertising are therefore non-excludable to producers of substitutable commodities that have not invested in advertising (Alston, Freebairn and James). As such, no individual producer has an incentive to advertise his product independently of other producers of close substitutes. Generic advertising therefore has some characteristics of a public good. Firms that have not invested in advertising campaigns may benefit as much from increased demand as those firms who do advertise; such is the typical free-rider problem. The result can be an underinvestment in commodity advertising (Alston, Freebairn and James).

3.2.2. Advertising Methods

Generic advertising is rarely undertaken at the firm level because the benefits are non-excludable. Rather, generic and commodity advertising is usually undertaken by an association of commodity producers. Examples of such associations are the International Wool Secretariat, the US Cattlemen's Beef Board and the California Milk Advisory Board. Producer associations in the US are administered by a board that is elected by producers, and operates under either the Federal Government's Secretary of Agriculture or the State Secretary. Similar programs exist in Canada. For example, the

Canadian Cattle Industry Development Council requires a $2 check-off per head of cattle.

Federal marketing orders originate from State Secretaries and compel all producers to contribute funds to their respective associations. Contributions are often made mandatory in an effort to avoid the free-rider problem outlined above, and to correct for the perceived underinvestment in advertising. Producer associations assert that by eliciting contributions from all producers, advertising campaigns can be undertaken that enhance the welfare of all producers.

Producer contributions are based on output so that those who stand to benefit most from an advertising campaign contribute a larger share of the budget. For example, beef producers contribute a fixed amount per head of cattle and milk producers contribute a fixed amount per volume of produced milk.

Contributions are spent on a variety of advertising media, including television, print and radio. A classic example of generic commodity advertising is “Pork, the other white meat.” This type of advertising aims to increase pork demand at the expense of pork substitutes, usually other meats (Alston, Freebairn and James).

3.3. Measuring Advertising Effectiveness

3.3.1. The Importance of Accurate Appraisal

There are several policy and welfare reasons for accurate appraisal of the effectiveness of advertising expenditures. An outline of the most important reasons follows. First, like any private firm that would evaluate the return to advertising expenditures, grower associations should likewise be interested in the success of their marketing efforts relative to the amount invested. Growers who contribute funds to association coffers should seek some assurance that their funds are being used effectively. The US$1000 (Hayes) that farmers contribute leads to higher total cost, and should be of interest to any profit-maximising producer. If production costs are rising, then the producer should expect to be compensated by higher prices. Furthermore, prices must rise sufficiently to more than offset the increase in production costs associated with the advertising program. That is, total revenue less total costs before advertising must be less than total revenue less total cost (including the advertising cost) after advertising. Otherwise there exists no motivation for a producer to contribute to a promotional campaign.

Producer associations have not been successful in convincing all of their members that check-off funds are in the best interest of all growers. There have been numerous cases (see Crespi for an overview of several such cases) in which producers were pursued in court by their respective associations for non-payment of mandated contributions. Some of the legal cases in the US have rested on the argument that forced contribution is a violation of a grower's First Constitutional Amendment right of free association -

that is, growers are being compelled to associate with their competitors via common advertising campaigns. However, it is likely that a producer belief that check-off programs are just not worth the cost underlies many of these cases. Another possibility is that non-compliant growers are simply trying to free-ride on their colleagues' promotion efforts. Either way, the effectiveness of advertising efforts is material to such cases.

The government, who must initially approve federal marketing orders, shares a common interest in measuring the success of marketing programs. However, the government's interest in advertising effectiveness goes beyond the desire to know if promotional campaigns increase demand for advertised products. Specifically, if increased demand for one product (beef) comes at the direct expense of decreased demand for a competing product (pork) then the government must be choosing one industry to support over another. This phenomenon is what Alston, Freebairn and James refer to as "beggar-thy-neighbour" advertising. Such a choice is sure to be contentious among producers of competing products.

Even if a government agency were able to justify supporting one group of producers at the expense of another, the potential for a negative or zero-sum result (if market growth due to advertising in one market just offsets market contraction in another) presents another problem. If the impacts of advertising just offset each other, then the costs of the advertising campaign generate a negative-sum result. Consider a meat industry with two products, beef and pork. If beef successfully increases demand through advertising and pork demand falls by a corresponding amount, then the net benefit to the meat industry is negative. Total demand in the industry is unchanged, however beef producers have contributed to a marketing program, thus increasing their costs. Furthermore, the administration of, and response to, advertising campaigns is sure to involve adjustment and transaction costs. If a campaign is successful in increasing demand, then the affected industry must increase production. Contrarily, the industry that faces decreased demand must decrease production. Non-zero adjustment costs result in negative net welfare effects for an otherwise zero-sum endeavour. Government involvement in a program that yields negative benefits is questionable in terms of net social welfare.

3.3.2. Techniques of Estimating Advertising Effectiveness

There exists a considerable literature that attempts to measure the effectiveness of commodity and generic advertising. The literature ranges from theoretical and technical attempts to advance the state of the art (see, for example, Baye, Jansen and Lee) to applied studies that focus on the returns to advertising in specific industries (for examples see Boetel and Liu or Piggott, et al.). This section outlines the most important and commonly used methods, and the following section critiques them in turn.

Perhaps the oldest and simplest method of evaluating consumer responses to advertising is market research. Producer groups survey consumers to assess the level of awareness for an existing advertising campaign. Such surveys may ask if consumers are aware of a campaign and ask whether responses are positive or negative. Market research is qualitative and does not provide a quantitative estimate of the effects of advertising on demand or profits.

One of the most common methods of estimating advertising effects is to estimate single-equation structural demand models. Some measure of demand (revenue, volume sales or consumption expenditures) is the dependent variable and exogenous variables such as income, demographic information and price are independent variables. To account for advertising, an independent variable that measures advertising expenditures is added to the equation. Advertising elasticities can be derived to quantify the effects of advertising on demand. Some examples of this methodology include Kinnucan, Chang and Venkateswaran and Lenz, Kaiser and Chung.

Kinnucan proposes a method to evaluate the effects of advertising on interrelated markets. Kinnucan develops a Muth-type disequilibrium model[34] that includes advertising as an exogenous variable. The model takes the form

$$d \ln Q_1 = -N_{11} d \ln P_1 + N_{12} d \ln P_2 + B_{11} d \ln A_1 \tag{3.1}$$

$$d \ln Q_2 = N_{21} d \ln P_1 - N_{22} d \ln P_2 - B_{21} d \ln A_1 \tag{3.2}$$

[34] Muth disequilibrium models consist of differential equations, often in demand, that illustrate relationships between changing prices, or other exogenous variables, and demand. See Muth (1965) for an exposition of the model.

Q_i, P_i and A_i represent quantities, prices and advertising expenditures for good i, and N_{ii} and B_{ii} are elasticities. The structural equations in (3.1) and (3.2) can be solved for reduced-form equations that provide estimates of how a change in advertising affects prices or quantities. Kinnucan applies his model to the US beef market using elasticities that were estimated in another empirical study.

The Kinnucan model's main insight is its recognition of the importance of product substitutability in determining the effectiveness of advertising. Specifically, a positive own-advertising elasticity (ie: B_{ii}) is no guarantee that advertising increases own price. If the differential equations in (3.1) and (3.2) are solved for $\ln P_1$, then the resulting sign is indeterminate and depends on the size of price, cross-advertising and supply elasticities. If, for example, B_{21} is large enough then, a fall in P_2 could be large enough to increase demand for good 2 and decrease demand for its substitute, good 1. The increase in P_1 initiated by increased advertising could be offset, or even reversed. Kinnucan describes this possibility as the result of *spillover*, or *feedback*, effects between substitutes.

Demand systems have become academic economists' favourite method for estimating the effects of advertising. The Rotterdam model (Thiel) and Deaton and Muellbauer's Almost Ideal Demand System (AIDS) are the most common. System methods allow demand for a group of separable goods to be estimated together, while accounting for substitution effects that Kinnucan outlines in his disequilibrium model. The specific properties and desirable attributes of system methods (specifically the AIDS) are expounded in chapter 3.5, however system methods share a common approach. Demand is modeled as a function of price, income, advertising expenditure and other exogenous variables. System estimation produces price and advertising elasticities to quantify the effects of advertising on all goods in the estimated system.

Demand systems have evolved to allow for dynamic adjustment processes. Anderson and Blundell's (1982) dynamic adjustment AIDS model includes, as an independent variable, a disequilibrium term that is measured by the difference between predicted demand and observed demand in the previous period. This type of adjustment framework theorises that short-run changes in demand respond to deviations of observed demand from predicted demand. The differences between dynamic adjustment AIDS models and the model developed in this essay are explained more fully in chapter 3.5.

A more recent approach to evaluate the effects of advertising uses time-series econometrics. Cavaliere and Tassinari test for long-run causality relationships between advertising and demand by means of a vector error correction (VEC) model. This model tests for and establishes the nonstationarity of price, demand and advertising data and estimates vector auto regression relationships (in error-correction form) between them. Variables are then tested for exogeneity as a means of determining each variable’s influence on other variables in the system. The primary contribution of this technique is to recognise the important time-series properties of the data used in estimation.

3.3.3. Problem with measurement methods

This section evaluates the measurement methods discussed in section 3.3.2 and concludes with some general measurement problems that apply to all techniques.

The primary benefit of market research is that investigators can formulate specific questions and pose them directly to consumers. Market research does not rely on aggregate market-level data and can provide detailed information about how consumers respond to specific advertising campaigns. Despite this advantage, market research is not a viable option for evaluating an advertising campaign on anything larger than a very local scale. Polling and questionnaires are costly, and acquiring a large enough representative sample for anything more than a local campaign would be a daunting and expensive task. However, if one assumes that consumers are rational, optimising agents, then market level data should suffice to explain consumer responses to advertising. If (granted, a big "if") comprehensive consumption, price and advertising data were available, then consumer responses to advertising could be estimated using a sound empirical model.

A more fundamental problem with market research is that surveys and questionnaires do not provide information that is required for policy decisions. Such research is qualitative in that it relates information about whether consumers respond to advertising. Policy decisions are enhanced by quantitative information on *how much* an advertising campaign affects consumer responses. Market research could be augmented to include contingent valuation, but the problems of cost and inadequate scope remain. Furthermore, market research is unable to control for the effects of other factors that may change demand. That is, *ceteris paribus* conditions cannot be enforced in a market research study.

Single equation structural demand models are appealing for two reasons. The first is that demand for each good can be modeled independently to include just those variables deemed necessary to determine demand for that good. The same group of independent variables need not be used for estimating demand functions for different goods (as is the case in some system methods of estimation). A second advantage is computational ease.

The advantages of single equation models are outweighed by the disadvantages. The primary disadvantage is that information is wasted which results in less efficient estimators (Greene). If a group of goods are related through substitution, then the error terms in each goods' demand function are likely to be related. For example, exogenous shocks that affect demand for one good may affect demand for closely related products. System methods can use this information to improve estimation. Furthermore, single-equation models are not consistent with consumer theory. Since single-equation demand models are not derived from an underlying utility or cost function, their specification is ad hoc and may be fragile (Piggott, et al.).

Kinnucan's disequilibrium model has the advantage of being simple, but it relies on elasticity estimates from other models. However, Kinnucan's article seems more an exposition of the problems associated with ignoring substitution effects than it does an attempt to estimate the market effects of advertising. Its primary contribution is to develop the analytical framework in which to recognise intra-market substitution effects.

System models possess several advantages. The first is that system methods allow the introduction of information variables without compromising the systems' theoretical integrity (see chapter 3.5 for an outline of the AIDS' theoretical properties). For example advertising expenditure, used as a proxy for advertising information, can be added to a demand system without affecting the parameter restrictions required to ensure consistency with consumer theory. Other variables such as health information (Boetel and Liu) or media information (Burton and Young) can be added in a similar fashion.

A second reason for analysing advertising within the context of a system is that estimation is done using system techniques. Such a setting is ideal for examining the cross-price and cross-advertising effects on substitutable goods within a weakly separable group. Estimating advertising responses using single-equation methods may waste information that is common to all equations. This information is more usefully applied in system estimation.

Hayes notes that estimating advertising effects in a system does impose a particularly strict constraint, however. Specifically, if the budget constraint is binding, then a zero-sum game is imposed. An advertising-induced increase in demand for one good must be offset by a corresponding decrease in demand for another good(s) in the estimated separable group. This restriction is illustrated in chapter 3.6. Another possible drawback of system estimation is that a misspecification in one equation is imposed on all equations in the system.[35] However, if one equation in a system is specified correctly, then it is likely that all are specified correctly. The independent variables in a demand function for, say, beef are likely to be the same as those in a demand function for chicken. This risk is likely outweighed by the statistical advantages of estimating demand equations in system form.

The chief benefit of the VEC analysis of advertising is that it handles properly the time-series characteristics of demand, price and advertising data. Most of these economic time series are non-stationary, and must be modelled accordingly. This issue is discussed in more detail in chapter 3.4.

Data problems are common across empirical economic studies, but there a few that are specific to advertising studies. First, MacDonald and Gould argue that by using advertising expenditures as an independent variable, the effectiveness of all advertising methods (i.e. print vs. television) are treated as homogeneous; such a treatment may generate biased estimates. If advertising markets are efficient, however, then relatively effective advertising methods are priced higher than relatively ineffective methods.

[35] System methods require that all equations contain the same exogenous variables.

Higher expenditure on advertising can be interpreted as either more purchases of less efficient methods, or fewer purchases of more efficient methods. As such, the effectiveness of advertising is implicit in its price and aggregating different types of advertising expenditures should not pose too large a problem.

Another data-related problem in advertising studies is the probable underreporting of advertising expenditure. Most studies utilise expenditures by grower or commodity associations as an independent variable. While such expenditure likely accounts for a large share of commodity advertising, it neglects all advertising that is funded by other sources. For example, advertising by a local commodity retailer (e.g. a grocery store advertising meat) would be excluded from the independent advertising expenditure variable. This amounts to a measurement error of the independent variable and results in biased parameter estimates. Such is the nature of advertising data, and one can only hope that reported data provides a good approximation of the actual data so that the estimation bias is limited.

Finally, all of the aforementioned methods (other than VEC analysis) neglect the time-series properties of the relevant data[36]. The consequences of ignoring the data's time-series properties are detailed in chapter 3.4. The importance of accurately modelling the long-run properties of the estimated data is particularly important in advertising studies. The time-series behaviour of prices in a demand system can provide important

[36] Dynamic adjustment AIDS models accommodate possible nonstationarity of the underlying data, but do not correctly model possible long-run relationships between variables. This point is addressed in greater detail in chapter 3.5.

insight into the effects of advertising in that market. This point is discussed in detail in chapter 3.4.

This essay develops a model that captures the beneficial aspects of the aforementioned methods, while attempting to minimise the drawbacks. The model is a demand system, so as to take advantage of desirable econometric properties and remain consistent with consumer theory. The model also considers the time-series properties of the data and models long-term relationships between variables. Doing so improves estimation and, more importantly, provides insight into the long-term dynamics in the estimated market.

3.4. The Importance of Dynamics and Time-Series Properties

3.4.1. Short-Run and Long-Run Effects

Static demand models provide parameter estimates that can be used to generate short-run elasticities. Such elasticities are frequently used to guide policy decisions on funding of commodity promotion programs. When making a policy decision, however, it is important to consider that the long-run effects of an advertising campaign may differ significantly from the short-run effects. Long-run demand and supply curves are generally flatter and more elastic than their short-run counterparts, and any comprehensive study of the effectiveness of an advertising campaign that involves long-term financial commitments from producers should consider long-run effects. A simple representation of the difference between short and long-run effects is illustrated in figure 3.2. Static supply and demand graphs are not entirely adequate tools for analysing the relevant dynamics, but some key points can be gleaned from graphical comparative statics. The initial equilibrium is characterised by price P^0 and the corresponding short and long-run demand and supply curves. If a producer association initiates an advertising campaign for their product, new costs are imposed on producers, shifting supply curves to retail up to S_{SR}^1 and S_{LR}^1. If the advertising campaign is successful in affecting consumer tastes, then demand curves shift up to D_{SR}^1 and D_{LR}^1. The new short-run equilibrium price is P^1. Over the long-term horizon, however, consumers have time to adjust their spending patterns and price competition from close substitutes pares away at demand for the promoted product. This is represented by a more elastic long-run demand curve. Long run price settles at P^2, where long-run demand and supply intersect.

The dynamics described above are a different phenomenon from advertising *wearout*. Advertising wearout (as described by Kinnucan, Chang and Venkateswaran) describes how advertising loses effectiveness over time because consumers become less responsive to promotional information. The response is psychological in nature, and is not related to price competition from close substitutes. Advertising wearout could affect a product with no substitutes. Consider a theoretical product that undergoes an advertising campaign that is successful in altering consumer tastes and shifting up short-run demand. If this campaign varies over time and is able to prevent advertising wearout, then consumer tastes do not change back to their pre-advertising state. Long-run demand shifts up along with short-run demand. If the promoted product has close substitutes, however, then price competition may pare away at demand. This is analogous to comparing a downward shifting short-run demand curve due to changing tastes to a downwards shifting short-run demand due to the fall in price of a substitute good. Wearout theory corresponds with the former and the effects of close substitutes the latter.

The effects of advertising wearout can be illustrated in figure 3.2. The initial advertising campaign alters consumer tastes and shifts short-run demand from D_{SR}^{0} to D_{SR}^{1}. Short-run equilibrium price is P^{1}. Note that long-run demand does not shift in the case of advertising wearout because the initial change in consumer tastes "wears off" over the long-term horizon. Long-run supply does, however, shift up to reflect the advertising levy imposed on producers. Once the effects of advertising wearout have taken hold and consumers have had the opportunity to adjust their spending patterns to account for cheaper substitutes, the long-run equilibrium price of P^{3} prevails at the intersection of D_{LR}^{0} and S_{LR}^{1}.

Supply-side market dynamics can also render advertising less effective over the long-term horizon. Refer again to figure 3.2, where the industry begins at price P^{0} and zero profits. An advertising campaign successfully shifts short-run demand up and leads to higher prices (P^{1}). Assuming low entry barriers, profit opportunity is a signal for firms to enter the industry. Firms enter over the long-term horizon, which is reflected by the relatively elastic long-run supply curve. Price falls back to P^{2} over the long-run horizon (or back to P^{3} if there is advertising wearout).

The important point to be drawn from this analysis is that the short-run effects of advertising may differ from the long-run effects; that is, short-run advertising elasticities may differ from long-run advertising elasticities. Any promotional campaign that requires continued funding from producers, such as a per-year check off system, should consider that the short-run impacts of the campaign may change as market dynamics unfold.

3.4.2. Time-Series Properties, Cointegration and Error Correction

Static demand models do not properly test for, or accommodate, the time-series properties of the data used in demand system estimation. The potential for nonstationary and cointegrated data series presents an additional challenge to the estimation of demand systems. This section outlines the importance of overcoming this challenge.

Most economic time-series data are non-stationary in levels. That is, the data are not characterized by a constant mean or variance over time. This is a violation of classical econometric assumptions, and modeling non-stationary data using classical econometric techniques can lead to spurious regression results (Granger and Newbold). Structural demand models that utilise such data must accommodate the time-series properties in order to avoid problems with parameter inference. Specifically, the t and F statistics obtained from a spurious regression tend to reject the null hypothesis of no relationship between variables too often, when in fact there is no meaningful relationship. A possible result of a spurious regression would be to attribute statistical significance to advertising, when there may exist no such significant relationship.

If data used in a demand system are nonstationary, then cointegration theory provides a method by which to analyse several non-stationary variables taken together. A group of variables may each be non-stationary, but they might share a non-stationary trend. It is possible that a linear combination of these non-stationary variables is stationary. This can be formally stated by the following: if $X_t = \left(X_{1t},..., X_{nt}\right)$ are non-stationary variables and integrated of order one, then there may exist up to $(n-1)$ $n \times 1$ vectors β such that $\beta X_t = \beta_1 X_{1t} + \beta_2 X_{2t} + ... + \beta_n X_{nt}$ is stationary. If so, then the variables in X_t are cointegrated of order (1,1). The cointegrating vectors β describe the variables' long-run relationships with each other.

A few important points are worth noting about cointegrated variables. First, a group of cointegrated variables need not move in a predictable manner. Rather, a linear combination of cointegrated variables should be stationary, even if each individual series is not. Second, cointegrated variables need not be equal to each other. One series can be permanently above (or below) all others, but such variables must be proportionally constant in the long run. Finally, cointegration does not mean that variables cannot deviate from their long-run equilibrium relationships. Just that when they do, market forces intervene in such a manner so as to return variables to their cointegrated equilibrium. Deviations can occur, but are temporary. This can be stated more formally by noting that cointegration does not require that $\beta X_t = 0$ in every period, just that $\beta X_t = e_t$, and e_t is stationary.

The establishment of cointegration between variables provides information that is valuable in estimating a structural relationship between variables. In fact, if a group of variables are cointegrated, then estimating a structural relationship without accounting for long-run dynamics amounts to a serious specification error (Enders). If a group of variables share a long-run equilibrium, then the variables' short-term dynamics should be, in part, influenced by the variables' relationships to their long-run equilibriums. Correcting for such a misspecification should improve model performance. Specifically, parameter estimates and elasticities may be more in line with theoretical expectations and will certainly be more econometrically sound. This is particularly relevant to demand systems, where counterintuitive results (positive own-price and negative own-advertising elasticities) are common.

Error-correction (EC) models allow structural and autoregressive models to accommodate the influence of a long-run equilibrium relationship on short-term dynamics. Once a cointegrating relationship between variables is estimated, then each period's deviation from the long-run equilibrium can be calculated. This deviation, called the EC term, can then be introduced into a regression equation that explains a variable's short-run dynamics. Intuitively, if a system of cointegrated variables is out of equilibrium in period *t*, then the system should respond in period *t+1* to move the system towards the long-run equilibrium.

A general error-correction model (ECM) takes the form:

$$\Delta Y_{it} = \sum_{k=1}^{r} \lambda_{ik} \mu_{kt-1} + \sum_{i=1}^{n} \sum_{j=1}^{\rho} \delta_{ij} \Delta X_{it-j} + \sum_{l=1}^{L} \gamma_{il} \Delta Z_{lt} + \varepsilon_{it} \text{ for } i = 1,\ldots,n. \qquad (3.3)$$

Y are dependent variables, *X* are contemporaneous independent and lagged dependent and independent variables. *Y* and *X* are cointegrated with each other and *Z* are exogenous (independent variables that are not cointegrated with *Y* and *X*) variables. *r* is the number of cointegrating vectors. μ_{kt-1} are lagged cointegrating vectors (or cointegrating terms) and λ_{ik}, δ_{ij} and γ_{il} are parameters. Δ is the difference operator.

Equation (3.3) is written so that both sides are stationary. The variables in *Y* are presumed to be integrated of order one, or I(1), so that the first difference of *Y* is I(0), or stationary. As such, the left-hand-side of (3.3) is stationary. The μ_{kt-1} terms are the residual errors from the cointegrating relationships between *Y* and *X*, and are therefore stationary. The variables in *X* and *Z* are also presumed to be I(1), so that their first-differences are stationary.

The model in (3.3) contains all stationary variables, as well as an EC term that accounts for the long-run dynamics of cointegrated variables. As such, (3.3) can be estimated using conventional techniques (least squares, seemingly unrelated regression, etc.) without concern of *spurious* regression. Equation (3.3) states that short-run changes in the dependent variable Y are functions of changes in other dependent variables, changes in independent variables and the size of the system's deviation from its long-run equilibrium in the previous period. If the system is out of equilibrium in period $(t-1)$, then changes in the dependent variable, Y, respond through the parameter λ.

3.4.3. Error-Correction and Agricultural Commodities

Demand systems are dynamic in nature. Advertising campaigns are unveiled over several periods, and consumers react with lagged responses. Market dynamics such as changing prices and shifting demand and supply curves unfold over time, making the problem of analysing advertising effectiveness a dynamic one. As such, the importance of investigating cointegrating relationships in advertising models goes beyond the econometric benefits of better statistical fits and corrections for misspecification errors. The existence of a long-run relationship between variables provides insight into the intra-market dynamics between substitutable products. This section outlines the importance and identification of these dynamics.

Advertising expenditures are made with the intention of providing consumers with information that changes their tastes, thereby shifting out the demand curve. The goal is a higher price and higher profitability. Shifting out the demand curve is contingent on the ability to differentiate one's product from close substitutes. Differentiating among a group of goods that includes, say, different brands of cars may be feasible. Even though such goods are often treated as a weakly separable group in demand systems, their characteristics differ enough that price is not the only differentiating factor. Certain groups of close substitutes, however, may not be so differentiable. Consumers may be readily willing to substitute one good for another based primarily on relative price changes. For example, consumers may choose to purchase pork instead of beef based on relative prices. Such nearly homogeneous goods are referred to as "closely substitutable commodities" (CSCs) through the remainder of this essay. The presumption is that the relative demand for one CSC over another CSC depends chiefly on relative prices. Beef and pork, for example, are likely to be CSCs.

If products can be categorised as CSCs then one would believe, *a priori*, that marketing programs intended to shift out the demand curve would be difficult propositions. Even if an advertising campaign were successful in increasing a single commodity's price in the short run by changing tastes and shifting out the short-run demand curve, market forces would act to bring its price back in line with those of its close substitutes over the long term. CSC prices, then, cannot diverge too far from a long-run equilibrium course before market forces intervene to return prices the equilibrated course of close substitutes. A short run demand and price effect may be observable (as illustrated in the model of figure 3.2), but as consumers alter their consumption patterns and substitute away from the now relatively higher priced CSC, there is downward pressure on demand and price for the advertised product.

The question facing policy makers is how to determine if a product can be categorised as a CSC. If it can, then it stands to reason that increasing profits by means of a marketing program may be difficult over the long-run horizon. As Gordon, Hannesson and Kerr point out, a preliminary test of whether a product is a CSC[37] may be worthwhile in evaluating an advertising campaign's probable success. Gordon, Hannesson and Kerr argue that CSCs are generally thought to be homogeneous and interchangeable, however a simple assertion that a product fits these characteristics may not be sufficient to warrant a decision about a potentially large-scale marketing campaign involving producer, processor and government funds. A testable definition of "CSC" is required.

Gordon, Hannesson and Kerr provide an insightful solution. If markets for multiple goods are related as CSCs, then there should exist a long-run equilibrium relationship between their prices. That is, prices for CSCs should be cointegrated with each other. If the price of a cointegrated CSC deviates from its long-run equilibrium relationship with other CSC prices, then market forces act to return that price to its equilibrium course.

It is possible that a promotional campaign successfully differentiates a promoted product from its substitutes, resulting in a permanent demand shift. For example, a crop may be found to have beneficial health attributes over and above its substitutes. Such information, if related to consumers, could render the product no longer a CSC - there are no close substitutes to pare away at the higher level of demand. Likewise, if the relevant industry is characterised by significant entry barriers, then the long-run supply

[37] Gordon, Hannesson and Kerr use the word "commodity" where this essay uses "CSC". Given the ambiguity about the definition of a commodity, CSC will be used throughout this essay.

may not be much more elastic than short-run supply. Price does not fall back as much in the long run.

It is also possible for a demand shock to break the cointegrated link between products that might have once been considered CSCs through, for example, discovery of a new use for an agricultural commodity (corn for ethanol production, for example). A new source of demand may emerge and price will not follow its once-cointegrated path. A preliminary test for cointegration among CSCs is, therefore, only useful from an historical perspective. That is, a finding of cointegration between prices tells us that the analysed products have shared a long-run relationship in the past.

There exists a natural synergy between the cointegration relationships described above and the type of advertising demand models that are frequently estimated for use in agricultural policy decisions. Most advertising demand models collect a group of goods together in a system and treat them as a weakly separable set of products. Though the assumption of separability is not sufficient to categorise a product as a CSC, testing for cointegration among prices for weakly separable goods (such as the groups of goods analysed in AIDS models) is particularly intuitive. As Rickertsen, Chalfant and Steen point out, the demand for a good within a weakly separable group depends only on the prices of the goods within the group and total expenditure on the group. It is reasonable to believe that if the price of one product in a separable group experiences a positive price shock, then consumers substitute away from that product and towards cheaper substitutes over the long-term horizon. The closer is the degree of substitutability, the larger is this effect. The error-correction term from equation (3.3) accounts for this effect in an empirical model. If prices are cointegrated and one price deviates from its cointegrated path, then the demand system's variables respond to the deviation through the parameter on the error-correction term.

An error-correction term is likely to play a particularly important role in estimating the differences between short-run and long-run elasticities. The closer is the degree of substitutability between products, the more consumers adapt their spending patterns over the long run in response to relative price differences. Short and long-run elasticities are therefore likely to differ more for CSCs than for products with no close substitutes. For this reason it is important to observe both short and long-run elasticities from a demand model that incorporates error-correction techniques when making policy decisions about closely substitutable agricultural commodities.

3.5. Model

This chapter describes the derivation of the traditional AIDS model and introduces a range of methods to incorporate advertising as an explanatory variable. The advertising-augmented AIDS model is then adapted to account for cointegrated variables by putting the model in error-correction form.

3.5.1. AIDS Model Derivation

Deaton and Muellbauer's AIDS is a researcher favourite in applied economics. The AIDS is used for all manner of applications, including computation of price and income elasticities, as well as estimating consumer responses to information variables that are introduced into the AIDS' underlying utility functions. It is the ability to introduce exogenous information variables, as well as its consistency with consumer optimisation behavioural assumptions, that have made the AIDS such a popular research tool. A brief outline of the AIDS derivation is presented below, and is from Deaton and Muellbauer's source article.

The AIDS model is derived from the PIGLOG class of cost functions[38], which defines the minimum expenditure required to attain a given level of utility at fixed prices. The cost function is defined as

$$\ln c(u, p) = (1-u)\ln\{a(p)\} + u\ln\{b(p)\}\ . \tag{3.4}$$

u lies between 0 (implying that the consumer is just achieving subsistence) and 1 (consumer is achieving bliss). Functional forms are assigned to $a(p)$ and $b(p)$, with enough parameters to allow the AIDS to be flexible. That is, enough parameters so that the cost function's derivatives can be set equal to an arbitrary cost function.

$$\ln\{a(p)\} = a_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2}\sum_k \sum_j \gamma^*{}_{kj} \ln p_k \ln p_j \tag{3.5}$$

[38] The PIGLOG class of cost functions are derived from "price independent generalised linear" budget share equations. Specific selection of the functional forms is required for practical application. See Deaton and Muellbauer for a discussion of these forms and of the aggregation properties of the PIGLOG cost functions.

and

$$\ln\{b(p)\} = \ln\{a(p)\} + \beta_0 \prod_k p_k^{\beta_k} \ . \tag{3.6}$$

These functions are inserted into (3.4) to obtain the AIDS cost function

$$\ln c(u,P) = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2} \sum_k \sum_j \gamma_{kj}^* \ln p_k \ln p_j + u\beta_0 \prod_k p_k^{\beta_k} \ . \tag{3.7}$$

Corresponding AIDS demand equations are obtained by Shephard's Lemma.

$$\frac{\partial \ln c(u,p)}{\partial \ln p_i} = \frac{p_i q_i}{c(u,p)} = w_i \tag{3.8}$$

where w_i represents good *i*'s budget share. The expenditure function in (3.7) can be solved for *u* and substituted into (3.8) to yield the AIDS budget share equations as functions of *P* and *X* (total expenditure).

$$w_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln\left\{\frac{X}{P}\right\}, \ i = 1, \ldots, n \tag{3.9}$$

where *P* is a price index, defined by

$$\ln P = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \ln p_k \ln p_j \tag{3.10}$$

and

$$\gamma_{ij} = \frac{1}{2}(\gamma^*_{ij} + \gamma^*_{ji}) \,. \tag{3.11}$$

Three groups of parameter restrictions are imposed when estimating (3.9) to ensure theoretical consistency.

$$\sum_{i=1}^{n} \alpha_i = 1 \quad \sum_{i=1}^{n} \gamma_{ij} = 0 \quad \sum_{i=1}^{n} \beta_i = 0 \,. \tag{3.12}$$

These restrictions ensure that budget shares sum to one (i.e. $\sum_{i=1}^{n} w_i = 1$).

$$\sum_{j=1}^{n} \gamma_{ij} = 0 \;\; . \tag{3.13}$$

Equation (3.13) ensures that the demand functions are homogenous of degree 0 (i.e. cost function is homogenous of degree 1).

$$\gamma_{ij} = \gamma_{ji} \;\; . \tag{3.14}$$

Equation (3.14) guarantees Slutsky symmetry of cross partial price derivatives.

The system in (3.9) can be estimated by either maximum likelihood or Seemingly Unrelated Regression method (Zellner).

The AIDS assumes a two-stage budgeting process in which consumers allocate income to a specified group of goods (the X in equation (3.9)) and then decide expenditures within that group. Demand for each good depends on the price of other goods in the group, and not prices of goods from other groups. Goods $i = 1, \ldots, n$ comprise a weakly separable group.

It should be noted that the AIDS, as outlined above, is nonlinear in parameters. The price coefficients in equation (3.10) interact with the real income coefficient (β_i) in equation (3.9), resulting in a nonlinear system. It is common practice to replace (3.10) with a linear approximation, defined as

$$\ln P^* = \sum_k w_k \ln p_k \tag{3.10a}$$

The linear, or Stone's, price index allows for linear estimation of the AIDS.

3.5.2. Advertising in the AIDS

Three methods of measuring advertising effectiveness in the AIDS are most popular among researchers. However, a few points about all methods are worth noting. All three of these methods use advertising expenditure as the independent variable. Also, each method can be linearised using Stone's price index in place of equation (3.10). The restrictions in (3.12), (3.13) and (3.14) can be imposed on the advertising-augmented systems to ensure theoretical integrity. A further restriction must be imposed on advertising coefficients to guarantee that shares sum to one. Also, lagged advertising variables can be inserted into all three of the forthcoming models. Lagged advertising variables act as proxies for consumers' accumulation of information over several periods.

The application of AIDS models to advertising has taken three primary forms (see Kinnucan, Thompson and Chang for a more detailed presentation of the following methods). In the first, advertising expenditures directly affect the share equation intercepts. The second method allows advertising expenditures to "deflate" prices throughout the share equations and the last method augments the price index of (3.10) with an advertising term. A brief outline of each methodology follows.

Modifying the share intercept

The first method allows advertising expenditures to affect a base level of consumption. Introducing advertising as an addition to the share equations' intercepts does this, while maintaining the AIDS' desirable properties. As such, advertising is presumed to act directly on consumer demand, and not indirectly through prices. This method of incorporating advertising is the most popular in applied studies - see Piggott, et al., Rickertsen, Chalfant and Steen, and Boetel and Liu for examples. The relevant consumer expenditure function is

$$\ln c(u,P) = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2}\sum_k \sum_j \gamma_{kj}^* \ln p_k \ln p_j + \sum_i \varphi_i \ln A_i + \sum_i \sum_j \phi_{ij} \ln A_i \ln A_j + \sum_i \sum_j \delta_{ij} \ln p_i \ln A_j + u\beta_0 \prod_k p_k^{\beta_k} \tag{3.15}$$

Shephard's lemma applied to (3.15) yields the estimable share equations

$$w_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln\left\{\frac{X}{P}\right\} + \sum_j \delta_{ij} \ln A_j \tag{3.16}$$

Advertising as a price deflator

This method allows advertising expenditures to act directly on consumers' perceptions of product prices. That is, advertising "deflates" observed prices, acting directly (through price terms) and indirectly (through the real income term) on the consumer's expenditure function. Weighted (by the δ_i term) advertising expenditures are subtracted from price terms in the cost function of equation (3.4) to give

$$\ln c(u,P) = \alpha_0 + \sum_k \alpha_k (\ln p_k - \delta_k \ln A_k) + \frac{1}{2}\sum_k \sum_j \gamma_{kj}^* (\ln p_k - \delta_k \ln A_k)(\ln p_j - \delta_j \ln A_j) + u\beta_0 \prod_k p_k^{\beta_k} \tag{3.17}$$

where A_k represents advertising expenditure on good *k*. Applying Shephard's lemma to (3.17) yields share equations

$$w_i = \alpha_i + \sum_j \gamma_{ij} (\ln p_j - \delta_j \ln A_j) + \beta_i \ln\left\{\frac{X}{P}\right\} \tag{3.18}$$

where the price index is now defined as

$$\begin{aligned} \ln P = \alpha_0 + \sum_i \alpha_i \left(\ln p_i - \delta_i \ln A_i\right) + \\ \frac{1}{2} \sum_j \sum_i \gamma_{ij} \left(\ln p_i - \delta_i \ln A_i\right) \cdot \left(\ln p_j - \delta_j \ln A_j\right) \end{aligned} \tag{3.19}$$

Advertising's deflationary effect can be seen by rewriting the price terms as

$$\ln\left(\frac{p_j}{A_j^{\delta_j}}\right). \tag{3.20}$$

An example of this method can be found in Green, Carman and McManus.

Advertising-augmented price index

A third alternative is to augment the AIDS price index with advertising terms. The price index in (3.10) can include advertising terms so that

$$\ln P = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \ln p_k \ln p_j + \sum_i \delta_i \ln A_i \tag{3.21}$$

The specification in (3.21) seems counterintuitive (Green, Carman and McManus). As advertising expenditure appears in the denominator of the real income term in the share equations, advertising can only have a positive effect on good *i*'s budget share if A_i and *lnP* are inversely related (ie: $\delta_i < 0$). Furthermore, β_i is expected to be negative for necessities. So even if A_i and *lnP* are inversely related (so that real income rises as advertising expenditures rise), advertising only has a positive effect for luxuries. See Green, Carman and McManus for an example of this methodology.

3.5.3. The Error-Correction AIDS Model

The prospect of a cointegrated equilibrium among prices begs the question of whether static demand-based studies of advertising effectiveness produce shortsighted conclusions. Consider a static AIDS model that estimates significantly positive advertising elasticities for a given product in a given spatial market. It seems possible that such studies are capturing the initial shift in demand that leads to a higher price (as illustrated in figure 3.2). The policy conclusion from such a study would be that advertising increases price. Policy makers in another spatial market considering a similar program for the same product would recommend pursuing a promotional campaign. Producers could also choose to continue an existing program because it is deemed to be successful in affecting demand.

If the market dynamics are as illustrated in figure 3.2 (i.e. the advertised good is a CSC), however, then the effect may be transitory. The price eventually returns to its cointegrated equilibrium course and profits are eroded. Since most promotional programs involve a long-term financing commitment from participants, costs may overtake benefits as market dynamics unfold. The speed at which these dynamics develop can be approximated by means of a VEC model that estimates speed-of-adjustment parameters. Speed of adjustment parameters provide an estimate of how much of a deviation from long-run equilibrium is corrected in each period. If the adjustment in prices to their long-run equilibrium is fast, then the costs of an advertising campaign may quickly overtake the benefits.

Industries that have undertaken promotional campaigns can be analysed using demand-based models that incorporate advertising expenditure. Estimated elasticities can quantify advertising's effect on consumption. Any such demand models should, however, be sure to investigate the time-series properties of the data and proceed accordingly. If the data are nonstationary, then an error-correction model should be pursued. An ECM produces more econometrically sound results by correcting for a misspecification error. Also, the long-run elasticities provide estimates of how consumption responds to an advertising shock after the long-run equilibrium is attained. This last point is key in the case of CSCs. Only long-run advertising elasticities relate how demand responds to advertising after the ‘dust has settled’ from the dynamics outlined in figure 3.2.[39] The AIDS specifications from section 3.5.2 can be enhanced as EC models. Each model from section 3.5.2 is expressed in error correction form below.

Advertising modifying the share intercept

$$\Delta w_{it} = \alpha_i + \pi_i \Delta w_{it-1} + \sum_j \gamma_{ij} \Delta \ln p_{jt} + \beta_i \Delta \ln\left(\frac{X}{P}\right) + \sum_k \delta_{ik} \Delta \ln A_{kt} + \lambda_i \mu_{t-1} \tag{3.22}$$

All variables in equation (3.22) are as previously defined, and are stationary in the EC form. There are a few key differences between the models in (3.22) and (3.16). First, all variables are first-differenced to account for nonstationarity. Second, a lagged dependent variable is included as an explanatory variable. It is through the parameter on the lagged dependent variable (π_i) that long-run elasticities are recovered. Third, the μ_{t-1} term is the lagged EC term which represents size of the deviation from the cointegrated variables' long-run equilibrium in period $(t-1)$. If the products in the demand system are CSCs and their prices are cointegrated, then they share a long-run equilibrium. If there exists a disequilibrium in period $(t-1)$, then the system's variables respond in period t to move back towards the long-run equilibrium. The change in w_i responds to this deviation according to the parameter λ_i. Note that if the products in the demand system are not CSCs and their prices are not cointegrated, then an error-correction term is not included in the model specification.

[39] Note that short-run elasticities are still useful in evaluating potential temporary benefits of advertising. This point is discussed further in chapter 3.8.

At this point, it is worth noting the difference between an EC-specified AIDS model and a dynamic adjustment AIDS model. A dynamic adjustment AIDS model (see Anderson and Blundell (1984) or Burton and Young for applications of such a model) appear similar to equation (3.22) but instead of the error correction term μ_{t-1}, the Anderson and Blundell (1982) version of a dynamic demand system includes the adjustment term $(\hat{w}_{it-1} - w_{it-1})$. $\hat{w}_{it-1}$ is the predicted value of demand share in period $(t-1)$, which is formulated by a static demand system. The current change in share, Δw_i, is a function of current exogenous variables as well as the disequilibrium between the predicted share in period $(t-1)$ and the observed share in period $(t-1)$.

Underlying the Anderson and Blundell (1982) model is the assumption that there exists a steady-state relationship that can be represented by the standard AIDS system of equation (3.9). Any deviation from this steady-state influences the movement of shares in future periods. Put another way, a long-run equilibrium relationship is presumed to exist between consumption shares and all right-hand-side variables (prices and real income). This equilibrium relationship is characterised by the parameters in equation (3.9).

There are several key difference between the Anderson and Blundell (1982) dynamic adjustment AIDS model and an EC-specified AIDS model. First, most dynamic adjustment models are restricted so that the adjustment coefficients are equal for all equations in the system (see Burton and Young). This is too restrictive an assumption, since each share may respond differently to a prior disequilibrium. A second important difference is that dynamic adjustment models do not test for and accommodate the time-series properties of the underlying data. Specifically, each variable must be tested for stationarity and treated accordingly. If some of the variables are nonstationary, then they must be tested for cointegration and modelled in EC form. A final difference between dynamic adjustment models and EC models is that dynamic adjustment models do not explicitly test for cointegrating relationships between nonstationary variables. The steady-state relationship of (3.9) is imposed and any deviation from that state is presumed to be a disequilibrium. Modern econometric techniques (which are described in chapter 3.4) allow for the explicit testing and estimation of cointegrating relationships. Those variables that are cointegrated should be included in a structural demand system as endogenous to the EC dynamic, and those that are not should be treated as exogenous.

The two other methods of incorporating advertising into a demand system can be transformed into EC form in a similar fashion, and the aforementioned comments apply to both.

Advertising as a price deflator

$$\Delta w_{it} = \pi_i \Delta w_{it-1} + \sum_j \gamma_{ij} \Delta(\ln p_{jt} - \delta_j \ln A_{jt}) + \beta_i \Delta \ln\left\{\frac{X}{P*}\right\} + \lambda_i \mu_{t-1} \qquad (3.23)$$

Advertising-augmented price index

$$\Delta w_i = \alpha_i + \pi_i \Delta w_{it-1} + \sum_j \gamma_{ij} \Delta \ln p_j + \beta_i \Delta \ln\left(\frac{X}{P}\right) + \lambda_i \mu_{t-1} \tag{3.24}$$

where

$$\ln P = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \ln p_k \ln p_j + \sum_i \delta_i \ln A_i \tag{3.25}$$

All of these models can, like their static counterparts, be estimated using Stone's price index as in equation (3.10a). The long and short-run elasticity derivations for these models are presented in the next chapter.

3.6. Application

This chapter applies the EC advertising-augmented AIDS model of the previous chapter to US meat (beef, poultry, pork and fish) data[40]. The meat industry is a logical choice for this type of study for several reasons. First, a large share of meat advertising is generic, and is initiated at the producer level. The Beef Industry Council and National Pork Producers Council in the US are responsible for promoting the interests of their member farmers and account for a large share of meat advertising. Since most meat marketing is done at the producer level, it would seem that producers view their products as homogenous. Such products might be categorised as CSCs.

[40] Other markets where the degree of product substitutability is even higher would be better suited to an EC AIDS model, however data is scarce. I had initially hoped to apply this model to the market for edible oils, in which there exists a very high degree of substitutability (i.e. corn oil vs. canola oil). The majority of advertising in the edible oils market is done at the firm level, however, and acquiring and aggregating such data is beyond the scope of this essay.

Another reason to analyse meats using this model is that supply is flexible. Several models that investigate the effectiveness of advertising are applied to supply-managed agricultural industries. Two common examples are eggs (Reberte, Schmit and Kaiser) and dairy (Lenz, Kaiser and Chung). There is a time lag for supply responses in meat as new stock grows, but output is not regulated. Profit opportunities can be followed by firm entry and increased supply over the long run.

The meat data for this study were graciously provided by Professor Brenda Boetel of The University of Wisconsin, River Falls. It comprises quarterly data, from 1976 to 1993. Price and consumption data are from Puttman and Allshouse and USDA's *Livestock and Poultry Situation and Outlook Report* (1976-2002). Fish consumption data are from the USDA's Economic Research Service, Food Consumption Data System. Advertising expenditure data were obtained from *AD $ Summary*, published by the Leading National Advertisers. Beef and pork advertising expenditures are those reported by the Beef Industry Council and the National Pork Producers Council.

3.6.1. Time-Series Properties of the Data

The first step in estimation is to establish the time series properties of the data. Specifically, all series must be tested for unit roots. If a unit root is found in any of the series, then the demand estimation strategy must reflect the data's nonstationarity.

The system variables are those outlined in equation (3.22): price, advertising expenditure, total group expenditure, consumption shares and real group expenditure. Advertising expenditures are deflated using the US Bureau of Labour Statistics Consumer Price Index. To account for the development of a stock of advertising knowledge and awareness, advertising expenditure is a three-quarter weighted average, with weights of 30-40-30 (Boetel and Liu). Consumption shares are calculated by multiplying price times quantity and dividing by group expenditure. This generates w_{it} variables for the demand equations. Real income is computed using Stone's price index from equation (3.10a) with share lagged one period. Note that all preliminary time-series testing (with the exception of consumption shares) is done in log form since the AIDS models are estimated in log form.

i) Unit Root Tests

The stationarity of the data is evaluated using the Augmented Dickey-Fuller (ADF) test. A preliminary step to unit root tests is to determine the optimal number of lags to include in subsequent ADF tests. The autocorrelation augmentation factor in the test is determined by estimating the equation

$$\Delta Y_t = a_0 + \gamma Y_{t-1} + a_2 t + \sum_{i=1}^{\rho} \beta_i \Delta Y_{t-i+1} + u_t \tag{3.26}$$

where ρ is selected by t-stat significance (i.e. the last lag with a significant t-stat is included in subsequent ADF testing). An alternative is to choose ρ based on minimum Schwarz Information Criterion.

The stepwise methodology developed in Enders is followed in performing ADF tests on the data. The non-standard Dickey-Fuller test statistic is needed only when deterministic regressors that are not in the actual data-generating process are included in the test equation (3.26). To ensure that the appropriate test statistic is used, Enders recommends the following procedure when the true data-generating process is unknown.

1. Test the null hypothesis of $\gamma = 0$ (i.e. series contains a unit root) in equation (3.26). The appropriate test statistic is τ_τ, from table A on page 439 of Enders. The least restrictive model includes a trend (t) and drift (a), and the test has a low power to reject the null hypothesis. If the null is rejected, then one can conclude that the series does not contain a unit root and testing is finished here.

2. If the null from step 1 is not rejected, then equation (3.26) must be evaluated to determine if it contains too many deterministic regressors (since too many regressors may have reduced the power of the test in step 1). To do this, the significance of the trend term is tested under the joint hypothesis $a_2 = \gamma = 0$ using the ϕ_3 statistic from table B on page 440 of Enders. If the trend is not significant, then proceed to step 3. If the trend is significant, then the presence of a unit root can be tested using the standard

normal distribution. If the null hypothesis of a unit root is not rejected, then one can conclude that the series does contain a unit root. If the null is rejected, then the series does not contain a unit root.

3. Estimate equation (3.26) without the trend and test for a unit root using the τ_μ statistic from table A on p. 439 of Enders. If the null of a unit root is rejected, then the series does not have a unit root. If the null is not rejected, then test the significance of the drift term in equation (3.26) using the joint test $a_0 = \gamma = 0$ and the ϕ_1 test statistic from table B on page 440 of Enders. If the drift is not significant, then proceed to step 4. If the drift is significant, then test for a unit root using the standard normal distribution. If the null of a unit root is rejected, then the series does not have a unit root. If the null is not rejected, then one can conclude that the series contains a unit root.

4. Estimate equation (3.26) without the trend and without the drift and test for a unit root using the τ statistic from table A on page 339 of Enders. If the null of a unit root is rejected, then the series has no unit root. If the null is not rejected, conclude that the series contains a unit root.

The aforementioned procedure is first applied to the data in levels. The stepwise procedure on all variables results in the conclusion that all series contain unit roots. The results are quite robust to selection of lag length in equation (3.26) and to significance level in null hypotheses testing. All modeled variables are, as expected, non-stationary in levels.

The next step is to confirm that all series are integrated of the same order. To do this, the same stepwise procedure described above is applied to all series in first-difference form. Testing on all series produces the conclusion that all series are stationary in first-differences. That is, all series are integrated of the first order. It should be noted that the testing of two of the series in first difference form is sensitive to the selection of lag length (ρ). Specifically, pork advertising expenditure and fish price unit root tests depend on lag length. However visual inspection of the series in first differences substantiates the conclusion that the data are stationary in first differences. The empirical investigation proceeds under the assumption that all data are integrated of order one.

Now that all series have been identified as being integrated of the same (first) order, long-run relationships between the series are investigated and estimated.

ii) Cointegration Tests

Cointegration between the nonstationary variables in the model is investigated using the Johansen procedure. The Johansen procedure is a multivariate generalisation of the Engle-Granger (Engle and Granger) test, which tests for stationary linear combinations of nonstationary variables. The methodology is briefly outlined below.

Consider the process

$$Y_t = a_1 Y_{t-1} + u_t \ . \tag{3.27}$$

Subtracting Y_{t-1} from both sides yields

$$\Delta Y_t = (a_1 - 1) Y_{t-1} + u_t \ . \tag{3.28}$$

If $(a_1 - 1) = 0$, then the process Y_t contains a unit root. This can be generalised to the multivariate case as

$$X_t = A_1 X_{t-1} + U_t \tag{3.29}$$

where X_t and U_t are $(n \times 1)$ vectors and A_1 is an $(n \times n)$ matrix of parameters. As is done in equation (3.28), subtract X_{t-1} from both sides to yield

$$\Delta X_t = (A_1 - I) X_{t-1} + U_t \tag{3.30}$$

where I is an $(n \times n)$ identity matrix, or

$$\Delta X_t = \pi X_{t-1} + U_t \tag{3.31}$$

where $\pi = (A_1 - I)$. If the rank of π is zero, then $\Delta X_t = U_t$, or $X_t = X_{t-1} + U_t$ and no linear combination of the variables in X is stationary. That is, the variables in X are not cointegrated. If, however, the rank of π is positive then there exists at least one linear combination of the variables in X that is stationary. So the rank of π is equal to the number of independent cointegrating vectors for the variables X.

This process can be generalised further to account for an autoregressive process.

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \ldots + A_\rho X_{t-\rho} + U_t. \tag{3.32}$$

Adding and subtracting $A_\rho X_{t-\rho+1}$ to (3.32) yields

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \ldots + (A_{\rho-1} + A_\rho) X_{t-\rho-1} - A_\rho \Delta X_{t-\rho+1} + U_t. \tag{3.33}$$

Successively continuing this process results in

$$\Delta X_t = \pi X_{t-1} + \sum_{i=1}^{\rho-1} \pi_i \Delta X_{t-i} + U_t \tag{3.34}$$

where $\pi = -(I - \sum_{i=1}^{\rho} A_i)$ and $\pi_i = -\sum_{j=i+1}^{\rho} A_j$. As in equation (3.31), the rank of matrix π is the number of independent cointegrating vectors. The Johansen procedure uses the fact that the rank of a matrix is equal to the number of characteristic roots that are not equal to zero. Characteristic roots are estimated, and then tested to evaluate how many are significantly different from zero. This provides the number of cointegrating vectors for the variables X.

The Johansen procedure is followed as outlined in Enders. Like the earlier ADF stepwise procedure, the Johansen methodology is followed in a manner than ascertains the best functional form for the estimation of equation (3.34).

The first step is to determine the optimal lag length of the vector autoregression (VAR) in equation (3.34). Optimal lag length is determined by minimising SIC criterion among VARs estimated from six lags to one lag. The SIC is minimised at one lag, and values are reported in table 3.1.

Determining the correct form of the deterministic regressors in equation (3.34) is the next step. Specifically, we test for the presence of an intercept in the cointegrating vector versus the alternative of an unrestricted drift term. This is done by estimating equation (3.34) with and without an intercept and ordering the set of characteristic roots of each equation ($\hat{\lambda}^*_1, \hat{\lambda}^*_2, ..., \hat{\lambda}^*_n$ and $\hat{\lambda}_1, \hat{\lambda}_2, ..., \hat{\lambda}_n$ respectively). If the unrestricted form (i.e. with no intercept) has r cointegrating vectors, then the statistic

$$-T \sum_{i=r+1}^{n} \left[\ln\left(1 - \hat{\lambda}^*_i\right) - \ln\left(1 - \hat{\lambda}_i\right) \right] \qquad (3.35)$$

has a chi-square distribution with $(n-r)$ degrees of freedom. The null hypothesis is the presence of an intercept in the cointegrating vector, and is rejected. The calculated test statistic is 41.33, which is greater than the chi-square statistic with 5 degrees of freedom at all levels of significance.

Now that the optimal number of lags in equation (3.34) and the form of deterministic regressors is determined, the cointegrating relationship is estimated. It should be noted that cointegration is investigated between price and consumption shares. Advertising expenditures are treated as exogenous.[41]

Equation (3.34) is estimated using the Johansen test procedure in Eviews. As determined in the previous testing, one lag is included and no intercept is included in the cointegrating equation. The Johansen estimation procedure produces characteristic roots, which are reported in table 3.2 along with the resulting trace and max test statistics.

The trace test statistic is calculated as

$$\lambda_{trace}(r) = -T\sum_{i=r+1}^{n}\ln\left(1-\hat{\lambda}_i\right) \tag{3.36}$$

[41] Degree of freedom restrictions prevent applying the Johansen procedure to all variables. Pair-wise Engle-Granger (Engle and Granger) tests show that advertising expenditures are not cointegrated with shares and prices.

and the null hypothesis is that there exist less than or equal to r cointegrating vectors versus a general alternative. If all $\hat{\lambda}_i$ are zero, then all λ_{trace} are zero. As $\hat{\lambda}_i$ get further from zero, λ_{trace} get larger and the null hypothesis is rejected more often.

The max statistic is calculated as

$$\lambda_{\max}(r, r+1) = -T \ln\left(1 - \hat{\lambda}_{r+1}\right). \tag{3.37}$$

The null hypothesis for the max test is that there exist r cointegrating vectors versus the specific alternative of $r+1$ cointegrating vectors. Like the trace test, the max test statistic grows larger as the characteristic roots diverge further from zero.

Both the trace and max tests concur that there exist three cointegrating vectors between the eight variables in the estimated system (four prices and four consumption shares). That is, there exist three different linear combinations of shares and prices that are stationary. These cointegrating vectors are used to calculate the EC terms in equation (3.3) for subsequent estimation of the EC AIDS model. The EC terms are operationalised by calculating a deviation from long-run equilibrium for each period using the coefficients from the estimated cointegrating vectors. Each period has three different EC terms, each having the form

$$\mu_t = a_1 \ln P_{bt} + a_2 \ln P_{ct} + a_3 \ln P_{ft} + a_4 \ln P_{pt} + a_5 w_{bt} + a_6 w_{ct} + a_7 w_{ft} + a_8 w_{pt} \tag{3.38}$$

where P_i and w_i are prices and shares for beef, poultry, fish and pork. One of the three cointegrating vectors, or error-correction terms, is selected for estimation in the EC AIDS model outlined below.

3.6.2. The Static and EC AIDS Models

The crux of this investigation is the estimation of the AIDS models outlined in chapter 3.5. The strategy is to estimate the model in static form, and to derive short-run advertising and price elasticities. The dynamic, EC counterpart is then estimated and its long-run elasticities are contrasted with short-run elasticities. It is difficult to form *ex ante* expectations about the relative sizes of the short versus long-run elasticities. Note that the derived elasticities are of quantity, not of price. Examination of the graphical model in figure 3.2 indicates that quantity changes could be ambiguous. This point is discussed further in chapter 3.7. Also regarding elasticities, it should be noted that there are several reasons to be cautious about using estimated elasticities in policy decisions. These reasons are discussed further in chapter 3.7.

Of the three techniques to include advertising in AIDS models that are discussed in chapter 3.5, only the first is presented in this essay. The second specification (advertising expenditure having a deflationary effect on prices) was estimated, but produced less appealing results. The third method of extending an AIDS model (advertising-augmented price index) is not pursued for the reasons discussed in chapter

3.5. All estimated models use Stone's price index as presented in equation (3.10a), but with share lagged one period.[42]

The model specification includes advertising as an addition to the share equations' intercepts. The estimated static form, which is referred to as model I, is

$$w_i = \alpha_i + \sum_{j=1}^{4} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{X}{P^*}\right) + \sum_{k=1}^{2} \delta_{ik} \ln A_k \tag{3.39}$$

for i = beef, poultry, fish and pork; k = beef and pork

where all variable are as defined in chapter 3.5. The system is estimated in Eviews by means of Zellner's seemingly unrelated regression (SUR) systems approach. One equation is dropped from estimation to avoid singularity of the covariance matrix since the equations sum to one by construction. The system is estimated by iterative SUR to ensure that parameter estimates are asymptotically invariant to the choice of the excluded equation. The equation for pork demand is dropped in all estimation hereafter. Pork equation parameters are recovered using cross-equation restrictions outlined in (3.12) through (3.14). There is an additional adding-up cross-equation restriction in advertising-augmented AIDS models. To ensure that consumption shares add up to one, the restriction

[42] Endogeneity of prices and advertising expenditure is not investigated in the current model. The focus of this research is how modeling demand in dynamic, error-correction form affects estimation results. Were this data set used for policy purposes, endogeneity tests may be worthwhile.

$$\sum_{i=1}^{n} \delta_{ij} = 0 \tag{3.40}$$

is imposed. It should be noted that imposing this restriction on advertising coefficients necessarily imposes the zero-sum game suggested by Hayes. An increase in demand for one good as the result of advertising must be met by a commensurate decrease in demand for another (or a combination of the others).

The EC counterpart to (3.39), which is referred to as model I_{EC}, is

$$\Delta w_i = \alpha_i + \pi_i \Delta w_{it-1} + \sum_{j=1}^{4} \gamma_{ij} \Delta \ln p_j + \beta_i \Delta \ln\left(\frac{X}{P^*}\right) + \sum_{k=1}^{2} \delta_{ik} \Delta \ln A_k + \lambda_i \mu_{t-1} \tag{3.41}$$

and is also estimated by iterative SUR. Parameter estimates, t-stats and R^2 are reported in table 3.3.

A few results from specifications I and I_{EC} are worthy of note. First, most own-price coefficients (γ_{ii}) are the expected (negative) sign. The exception is fish price in the static representation of model I's fish share equation. However it should be noted that the estimated coefficients are not direct reflections of computed elasticities. In fact, the sign of a coefficient may not match the sign of its respective elasticity. Cross-price coefficient estimates (γ_{ij}) show mixed results. Some are the expected positive sign, while some are negative. AIDS models are replete with mixed results on the signs and significance of price parameter estimates; these results are not unusual.

Another interesting point is the change in the significance of advertising coefficients (δ_{ij}) between model I and model I_{EC}. The t-stats for all advertising coefficients fall substantially in the EC model. This result suggests that the statistical significance of advertising expenditure's effect on consumption shares is smaller when modelled in dynamic EC form than in static form. Correcting for the misspecification error in the static model results in less significance being attributed to advertising as a determinant of demand.

A third interesting result is that all coefficients of adjustment (λ_i) are negative in the EC form of model I_{EC} (with the exception of fish). This concurs with expectations, and implies that a deviation from the cointegrated long-run equilibrium in period *t* is partially corrected in period *(t+1)*. A positive (negative) error reflected in the μ_{t-1} variable has a negative (positive) effect on the relevant consumption share in the next period. This dynamic works to move the system back towards its long-run equilibrium.

3.6.3. Elasticities

This section derives advertising elasticities for the estimated models. Note that the Green and Alston methodology is not necessary in this case because lagged shares are used in the formulation of Stone's price index. However, as an additional resource, the derivation of advertising elasticities in the case of Stone's price index with contemporaneous share is included in the appendix.

Short-run price elasticities are derived using the following method. Define

$\eta_{ij} = \dfrac{d \ln q_i}{d \ln p_j}$ as the elasticity of demand for good *i* with respect to price of good *j*. To

derive the formula for this elasticity, first note that

$$w_i = \frac{p_i q_i}{c(u, p)} \tag{3.42}$$

or

$$q_i = \frac{c(u, p)}{p_i} w_i \tag{3.43}$$

so that

$$
\begin{aligned}
\eta_{ij} &= \frac{dq_i}{d\ln p_j}\frac{1}{q_i} \\
&= \frac{1}{q_i}\frac{d\left(\frac{c(u,p)}{p_i}w_i\right)}{dp_j} \\
&= \frac{1}{q_i}\left[\frac{c(u,p)}{p_i}\frac{dw_i}{d\ln p_j} + w_i c(u,p) p_j \left(\frac{d\left(\frac{1}{p_i}\right)}{dp_j}\right)\right] \\
&= \frac{1}{w_i}\frac{dw_i}{d\ln p_j} - \frac{1}{q_i}\left[w_i c(u,p) p_j \left(\frac{d\left(\frac{1}{p_i}\right)}{dp_j}\right)\right] \\
&= \frac{1}{w_i}\frac{dw_i}{d\ln p_j} - \partial
\end{aligned}
\tag{3.44}
$$

where ∂ is the Kronecker delta.[43] Short-run advertising elasticities are derived using the same technique and are calculated as

$$
\varepsilon_{ij} = \frac{1}{w_i}\frac{dw_i}{d\ln A_j}. \tag{3.45}
$$

Long-run elasticities for model I_{EC} are calculated by dividing the computed short-run elasticities by $(1-\pi_{ii})$. The long-run advertising elasticity is defined as

$$
\varepsilon_{ij}^{LR} = \frac{\varepsilon_{ij}}{(1-\pi_{ii})}. \tag{3.46}
$$

[43] The Kronecker delta equals one if $i = j$ and zero otherwise.

Elasticities are reported in table 3.4 and all are uncompensated. Before discussing the estimated elasticities, a discussion of how to interpret each elasticity is warranted. Model I generates only short-run elasticities. These elasticities reflect the effect on quantity demanded of advertising-induced short-run demand and short-run supply shifts. This initial movement is characterised in figure 3.2 by the movement of the point at P^0 to that at P^1. The elasticities from model I are estimated without including the error-correction term, however. The existence of a long-run relationship between variables in the estimated system tells us that the error-correction term is an important factor in determining short-run changes in dependent variables. Omission of this factor results in specification error.

Specification I_{EC} also generates short-run elasticities, but after correcting for the specification error. As such, the elasticity estimates from the error-correction specification should be more econometrically and theoretically sound.

Long-run elasticities are derived from model I_{EC}. The π_{ii} parameter is used to derive the long-term effects through equation (3.46). Long-run elasticities reflect how quantity demanded responds to advertising expenditure once the long-run equilibrium has been attained. This is at the intersection of the long-run demand and long-run supply curves in figure 3.2.

The task now is to compare the short-run elasticities from model I to model I_{EC} and to compare short-run elasticities to long-run elasticities. A comparison of short-run elasticities from model I to model I_{EC} illustrates how correcting for the static model's misspecification error affects model results. A comparison of short-run to long-run elasticities is important to differentiate short-run from long-run advertising effects.

The own-price elasticities from model I are all negative, thereby concurring with a priori expectations. Some cross-price elasticities are negative, implying complimentarity between meats. The magnitudes are reasonable within the context of other meat demand studies (including Boetel and Liu, Piggott, et al., Burton and Young) but it should be noted that making such comparisons is a questionable endeavour.[44]

There are a few noteworthy differences between the short-run elasticities from model I and those from model I_{EC}. First, most own-price elasticities are smaller in the EC model (with the exception of fish). The inclusion of the EC term in model I_{EC} gives the impression of less price elastic meat demand. Also, most cross-price elasticities are smaller in model I_{EC} than in model I. The exception remains fish. Smaller elasticities are a result of correcting for the static model's misspecification error. The inclusion of the error-correction term reduces the influence that prices have on the movement in shares. The system's short-run movements towards its long-run equilibrium, as reflected in the error-correction term, controls for some of the shares' short-run deviations, thus resulting in smaller price elasticities.

[44] Different functional forms, model specification and data sets frequently provide markedly different estimates. It is hoped that the model proposed in this essay is one step towards a methodology that provides more robust elasticity estimates.

Long-run price elasticities from model I_{EC} do not provide support for the proposition that demand is more price elastic in the long-run than in the short-run. Pork is the only meat that has a larger own-price elasticity in the long-run than in the short-run. This result is counterintuitive and highlights the caution that must be exercised in interpreting elasticity estimates from demand systems. Unfortunately, the static model does not provide long-run elasticities which could be compared to the long-run elasticities from the error-correction model. As such, it cannot be determined if modeling in EC form provides an improvement in the estimation of long-run price elasticities. The use of elasticity estimates in policy decisions is discussed further in chapter 3.7.

Advertising elasticity estimates from models I and I_{EC} make for an interesting comparison. The short-run own-advertising elasticity for beef in model I is negative, implying that beef advertising has a negative impact on beef demand. This is an intuitively unappealing result. Note, however that short-run own-advertising elasticity for beef is positive in model I_{EC}; this is more in line with theoretical expectations. Another key difference between short-run advertising elasticities in model I and model I_{EC} is that most cross-advertising elasticities become negative when modelled in EC form. Correcting for the misspecification error by including an EC term generates the more intuitively appealing result that advertising for one type of meat has a negative impact on demand for other meats.

Long-run advertising elasticities are similar to their short-run counterparts, however most are smaller in magnitude. This result suggests that consumer demand is less responsive to advertising over the long-term horizon. The exception is pork, for which long-run advertising elasticity is larger than short-run advertising elasticity. The larger long-run pork advertising elasticity suggests that pork producers have been successful in attempts to sustain, and even magnify, the effects of their advertising messages. This could be attributable to an accumulation of information, as discussed in chapter 3.5.2.

The broad conclusions from elasticity estimates are two-fold. The first is that the EC model of I_{EC} generates elasticity estimates that are more in line with theoretical expectations, especially in the case of advertising. All own-advertising elasticities are positive and most cross-advertising elasticities are negative. The second conclusion is that there are differences between short and long-run elasticities. Whether elasticities become smaller or larger depends on the meat type, but producers should be aware that the instantaneous effects of an advertising campaign may not be long lasting.

3.7. Demand Models and Elasticities in Policy Decisions

Elasticities are the common yardstick for determining if advertising "works". Demand systems are estimated, elasticities computed and policy decisions made according to those elasticities. There are several reasons, however, to exercise caution when evaluating elasticities (see Alston and Chalfant for a thorough examination of the sensitivity of estimation results to model specification).

First, as in any empirical economic study, the *ceteris paribus* conditions are difficult to enforce. One would hope that a properly specified demand equation controls for all relevant variables that affect demand. This is never the case. Misspecification affects parameter estimates and attributes either too much or too little influence to some or all of the independent variables in the system. For example, an exogenous shock that increases beef demand and coincides with an advertising campaign for beef attributes too much credit to advertising if the exogenous shock is not in the demand system. Demand systems already contain several variables, and the benefits of including more exogenous variables must be balanced with the benefits of econometric parsimony.

Elasticities are based on these imperfect parameter estimates and must be understood in this context. Elasticities can be considered, at best, crude estimates of the direction and magnitude of an independent variable's effect on a dependent variable. As such, one must avoid deriving overly-fervent conclusions. For example, Chang and Green estimate negative own-advertising elasticities for dairy products. Based on these elasticities, they state that "consumers respond...negatively to advertising for dairy products...". This statement implies that consumers view dairy advertisements and consciously decide to decrease their consumption of dairy products in response. This seems implausible. It is much more likely that their model is not picking up the effects of some other factor that determines demand for dairy products. That is, the presumed *ceteris paribus* is not so *paribus*. Chang and Green likely understand this problem, however such statements must be interpreted cautiously.

A second consideration is that the elasticities derived in AIDS models are elasticities of *quantity*, not of *price* or of *expenditure*. That is, if advertising elasticity is positive and significant, then one can conclude that advertising increases quantity demanded of the advertised product. That is, q_i increases. The supply-side dynamics described in figure 3.2 demonstrate that advertising could lead to increased market consumption, with no significant effect on price. All that advertising accomplishes in this case is a larger industry. Profits do not rise if the industry faces constant or rising costs. An intermediate case where both price and quantity increase is also a possibility, but the profit increase is not as large as if only price increases.

Furthermore, even if price does increase as a result of advertising, at least some portion of that price increase must be attributed to higher producer costs. Figure 3.2 illustrates how supply shifts left when producers are faced with new advertising costs, often in the form of a per-unit check off. A price increase of this sort does not increase profits.

Finally, the size of advertising elasticities must not be equated with the size of potential returns for producers (Green, Carman and McManus). Even if the actual advertising elasticity is positive and significant, producer returns depend on factors such as the cost of expanding production to meet higher demand and the cost of the advertising campaign.

3.8. Conclusions

The debate about the effectiveness of agricultural advertising campaigns is likely to remain active. Policy makers and producers have vested interests in knowing the usefulness of programs to which they are compelled to contribute funds. The job of analysing the effectiveness falls on the shoulders of economist, who must use aggregated data to estimate the effects of advertising on demand. Such estimation usually takes the form of demand systems.

Demand systems are useful in that they provide quantitative estimates of how advertising expenditure changes demand. The agricultural economics literature is replete with such studies. Results from these studies are often unsatisfying, and are frequently sensitive to model specification. Also, the majority of demand system studies are static and therefore consider only short-run effects. Producers who are compelled to contribute to a marketing program over a long-run horizon should be interested in the long-term prospects of success in promoting their product.

This essay attempts to improve the state of assessing advertising effectiveness in two primary ways. The first is to outline the importance of long-run versus short-run effects in analysing advertising effectiveness. The second is to stress the importance of correctly modelling the time-series properties of the data used in demand system estimation. Correctly accounting for these properties should produce a more econometrically sound, and hopefully robust, method of assessing advertising. This essay also explains the particular relevance of correctly modelling time-series properties and accounting for long-run effects in the context of agricultural commodities.

An empirical model is developed that accounts for the time-series properties of the data and provides long-run elasticity estimates. The error-correction AIDS model is applied to US meat data and results are compared to a traditional static AIDS model. There are important differences between the traditional static demand model and the EC model that is developed in this essay; the estimated parameters and elasticities show that accounting for time-series properties and including an error-correction term does produce different results than traditional static models. Analysis of the data reveals a long-run relationship between meat prices in the US, indicating that these products may fit into the closely-substitutable commodity categorisation. This long-run relationship is estimated in a vector error-correction model and incorporated into a demand system by means of error correction.

The error-correction model corrects for the misspecification in the static demand model, and produces markedly different empirical results. Short-run elasticities from the error-correction model differ from elasticities derived from the static model. Also, long-run elasticities are different from short-run elasticities. This suggests that the long-run effects of an advertising campaign may be different than the short-run effects. The error-correction model also produces some more intuitively appealing results.

A comprehensive appraisal of a commodity advertising program would require more information that can be gleaned from a demand system. Specifically, the estimation of a vector error correction model could provide information about how quickly, and for how long, prices are likely to respond to an advertising-induced price shock. Short-run elasticities would quantify this response. Also, the costs an advertising program and the costs of increasing output to meet higher demand would have to included in any benefit-cost study.

It is important to remember that the tools used by economists to estimate policy effects have limitations and are far from perfect. With that proviso in mind, it is likely that demand models will continue to be used in policy decisions about agricultural commodity promotion. Economists should strive to make these demand models as reliable and comprehensive as possible.

Figure 3.1. Advertising-Induced Demand Shift

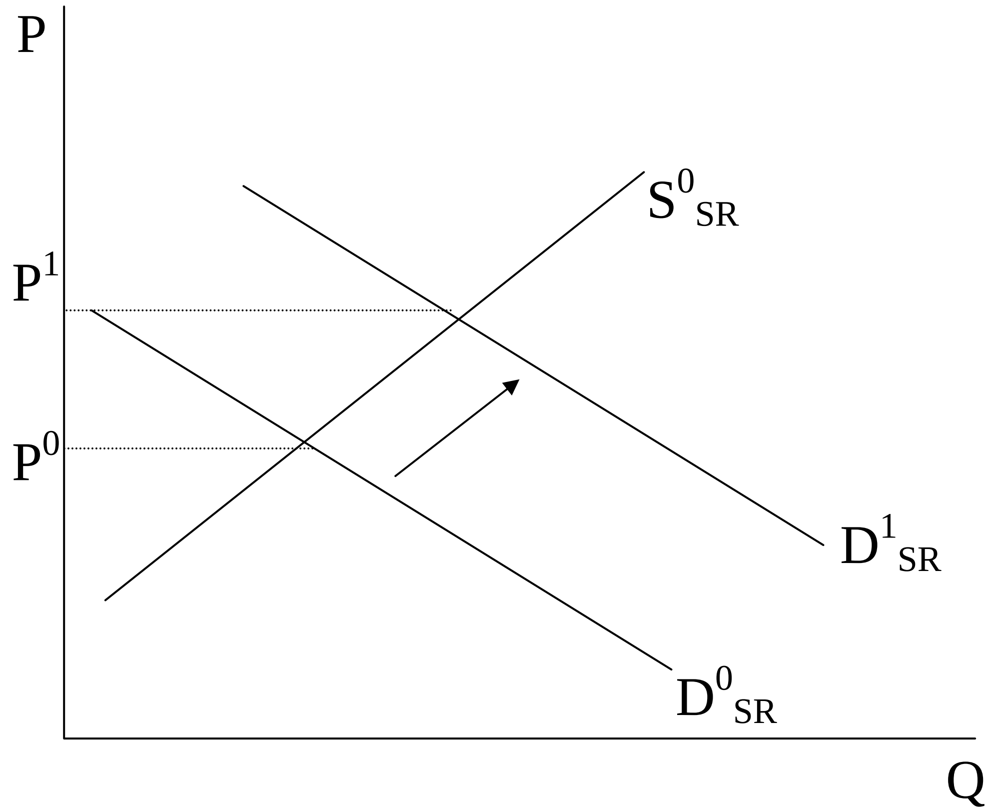


Figure 3.2. Market Force Dynamics

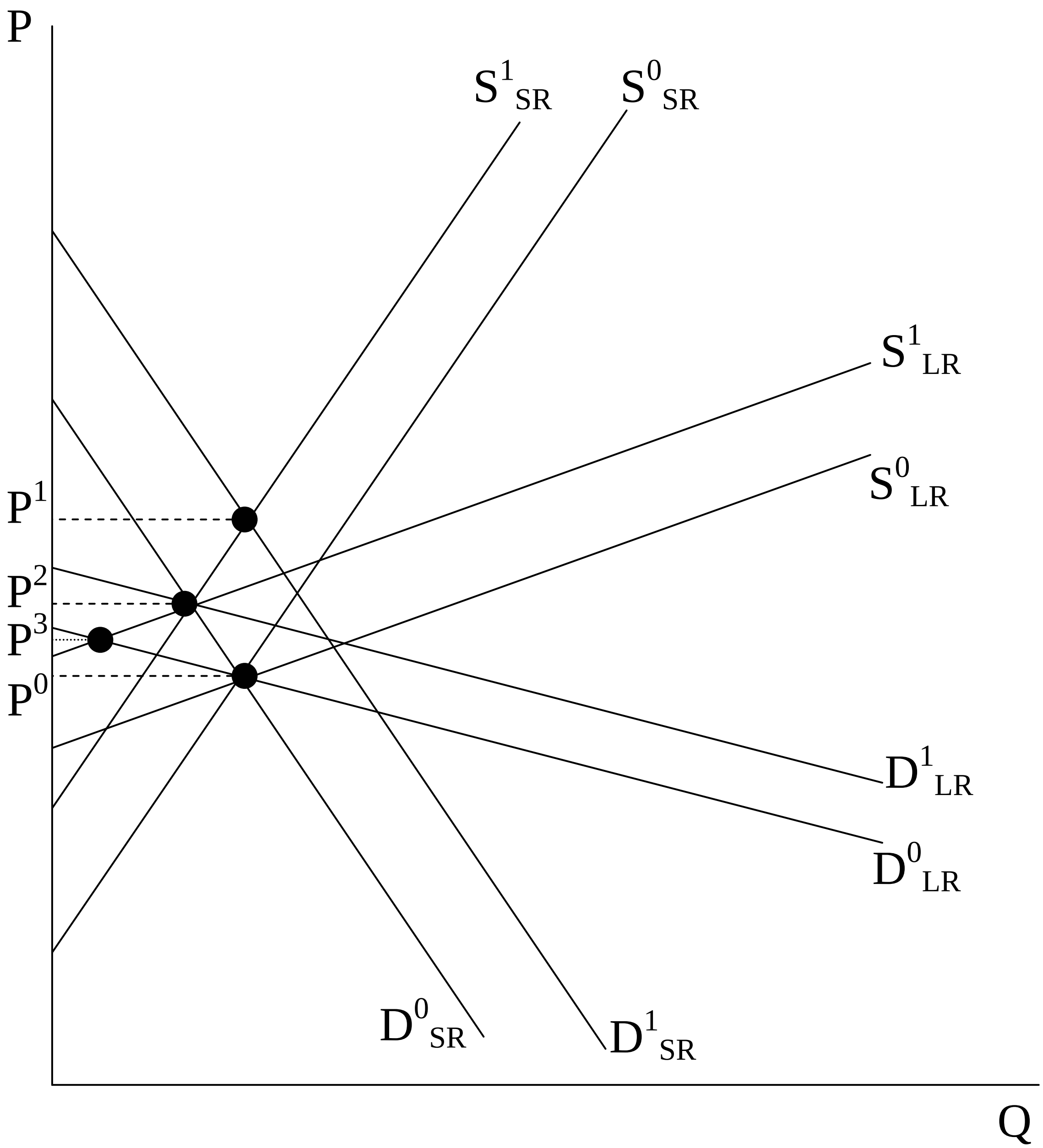


Table 3.1. VAR Tests

Lags	SIC
1	-47.96279
2	-45.93485
3	-44.10836
4	-43.54338
5	-42.92391
6	-43.24723

Table 3.2. Johansen Tests

Max test

Hypothesis	Eigenvalue	Test Statistic	5% Critical Value	Conclusion
r=0	0.6623	75.9827	47.99	Reject
r=1	0.5726	59.4977	41.51	Reject
r=2	0.4456	41.2884	36.36	Reject
r=3	0.2956	24.5298	30.04	Not Reject
r=4	0.1670	12.7940	23.80	Not Reject
r=5	0.1002	7.3894	17.89	Not Reject
r=6	0.0482	3.4552	11.44	Not Reject
r=7	0.0026	0.1833	3.84	Not Reject

Trace test

Hypothesis				
r=0	0.6623	225.1206	141.21	Reject
r < or = 1	0.5726	149.1379	109.99	Reject
r < or = 2	0.4456	89.6402	82.49	Reject
r < or = 3	0.2956	48.3518	59.46	Not Reject
r < or = 4	0.1670	23.8220	39.89	Not Reject
r < or = 5	0.1002	11.0279	24.31	Not Reject
r < or = 6	0.0482	3.6385	12.53	Not Reject
r < or = 7	0.0026	0.1833	3.84	Not Reject

Table 3.3. Parameter Estimates

	I	IEC		I	IEC
α_b	0.7036	-0.0014	α_f	0.0708	-0.0017
	32.7525	-1.0865		7.8904	-1.2074
π_b		-0.1886	π_f		-0.4084
		-1.7227			-3.5192
γ_{bb}	-0.1566	-0.0276	γ_{fb}	-0.0054	0.0364
γ_{bc}	0.0739	0.0034	γ_{fc}	-0.0001	0.0164
	3.2349	0.3308			
γ_{bf}	-0.0054	0.0364	γ_{ff}	0.0094	-0.0843
	-0.4407	1.8499			
γ_{bp}	0.0881	-0.0123	γ_{fp}	-0.0039	0.0315
	7.9826	-0.8516		-0.4658	1.8983
β_b	0.0083	-0.0177	β_f	0.0177	0.0753
	0.2491	-0.8179		1.2727	2.5932
λ_b		-0.0039	λ_f		0.0072
		-1.3370			2.1974
δ_{bb}	-0.0068	0.0033	δ_{fb}	0.0022	0.0015
	-3.1602	1.8789		2.0594	0.8091
δ_{bp}	-0.0222	-0.0018	δ_{fp}	0.0025	-0.0023
	-5.5181	-0.9089		1.3426	-1.0692
R^2	0.9238	0.1495	R^2	0.6087	0.3620

	I	IEC		I	IEC
α_c	0.0488	0.0023	α_p	0.1769	0.0008
	2.3502	2.5492			
π_c		-0.0097	π_p		0.6068
		-0.0767			
γ_{cb}	0.0739	0.0034	γ_{pb}	0.0881	-0.0123
γ_{cc}	-0.0662	-0.0124	γ_{pc}	-0.0076	-0.0075
γ_{cf}	-0.0001	0.0164	γ_{pf}	-0.0039	0.0315
	-0.0098	1.4296			
γ_{cp}	-0.0076	-0.0075	γ_{pp}	-0.0767	-0.0117
	-0.7234	-0.7649			
β_c	0.0253	-0.0277	β_p	-0.0512	-0.0299
	0.8200	-1.8995			
λ_c		-0.0019	λ_p		-0.0015
		-0.9439			
δ_{cb}	0.0098	0.0005	δ_{pb}	-0.0052	-0.0054
	4.9518	0.4614			
δ_{cp}	0.0161	-0.0002	δ_{pp}	0.0036	0.0043
	4.1955	-0.1377			
R^2	0.8911	0.0803			

Table 3.4. Elasticities

Uncompensated Price Elasticities

Short-run elasticities for Model I

	Beef	Poultry	Fish	Pork
Beef	-1.3989	0.1789	-0.0151	0.2145
Poultry	0.2452	-1.2801	-0.0081	-0.0543
Fish	-0.1570	-0.0589	-0.8995	-0.1067
Pork	0.4188	0.0222	0.0007	-1.2443

Short-run elasticities for I_{EC}

	Beef	Poultry	Fish	Pork
Beef	-1.0511	0.0201	0.0944	-0.0192
Poultry	0.0560	-1.0200	0.0717	-0.0011
Fish	0.0784	-0.0395	-2.1359	0.1500
Pork	-0.0474	0.0011	0.1305	-1.0151

Long-run elasticities for I_{EC}

	Beef	Poultry	Fish	Pork
Beef	-0.8843	0.0169	0.0794	-0.0162
Poultry	0.0555	-1.0102	0.0710	-0.0011
Fish	0.0557	-0.0281	-1.5165	0.1065
Pork	-0.1206	0.0029	0.3318	-2.5816

Advertising Elasticities

Short-run elasticities for Model I

	Beef	Pork
Beef	-0.0171	-0.0553
Poultry	0.0378	0.0619
Fish	0.0273	0.0320
Pork	-0.0199	0.0137

Short-run elasticities for I_{EC}

	Beef	Pork
Beef	0.0083	-0.0045
Poultry	0.0021	-0.0007
Fish	0.0195	-0.0293
Pork	-0.0208	0.0166

Long-run elasticities for I_{EC}

	Beef	Pork
Beef	0.0070	-0.0038
Poultry	0.0021	-0.0007
Fish	0.0138	-0.0208
Pork	-0.0530	0.0423

APPENDIX - DERIVATION OF ADVERTISING ELASTICITIES WHEN USING STONE'S PRICE INDEX

If an AIDS model is estimated using Stone's price index and share is not lagged in equation (3.10a), then the elasticity derivations are not as shown in chapter 3.6. Using standard AIDS elasticity formulae is a common error in empirical demand analysis that utilises Stone's price index (Green and Alston). Green and Alston's article outlines how to solve for price elasticities, and this appendix illustrates how to derive the correct advertising elasticity formulas.

Stone's price index allows for the linearization of the standard AIDS model, however elasticity derivations are made considerably more complicated by the inclusion of lagged consumption shares on the right-hand-side of the estimated equation. Specifically, when taking the derivative of the price index, the partial derivatives of all consumption shares (see equation (3.10a)) must be computed. Each elasticity is a function of itself and all other elasticities. This problem is outlined and then solved below.

Define $\varepsilon_{ij} = \dfrac{d \ln q_i}{d \ln A_j}$ as the elasticity of demand for good *i* with respect to advertising expenditure on good *j*. To derive the formula for this elasticity, first note that

$$w_i = \frac{p_i q_i}{c(u, p)} \tag{3.A.1}$$

or

$$q_i = \frac{c(u, p)}{p_i} w_i \tag{3.A.2}$$

so that

$$\begin{aligned} \varepsilon_{ij} &= \frac{dq_i}{d \ln A_j} \frac{1}{q_i} \\ &= \frac{1}{q_i} \frac{c(u, p)}{p_i} \frac{dw_i}{d \ln A_j} \\ &= \frac{1}{w_i} \frac{dw_i}{d \ln A_j} \\ &= \frac{d \ln w_i}{d \ln A_j} \end{aligned} \tag{3.A.3}$$

In the case of specification I, this is

$$\varepsilon_{ij} = \frac{1}{w_i}\left(\delta_{ij} - \beta_i \frac{d \ln P^*}{d \ln A_j}\right) \tag{3.A.4}$$

The derivative of the price index in the standard AIDS model is a linear function of parameters. The derivative of Stone's price index is more complicated.

$$\frac{d \ln P^*}{d \ln A_j} = \sum_k \ln P_k \frac{dw_k}{d \ln A_j} = \sum_k w_k \ln p_k \frac{d \ln w_k}{d \ln A_j} \tag{3.A.5}$$

To solve for the $\frac{d \ln w_k}{d \ln A_j}$ term, relationship (3.A.3) shows that

$$\varepsilon_{ij} = \frac{d \ln w_i}{d \ln A_j} \tag{3.A.6}$$

Substituting this last result into (3.A.5) yields

$$\frac{d \ln P^*}{d \ln A_j} = \sum_k w_k \ln p_k \varepsilon_{kj} \tag{3.A.7}$$

Plugging this into (3.A.4) gives

$$\begin{aligned} \varepsilon_{ij} &= \frac{1}{w_i}\left(\delta_{ij} - \beta_i \sum_k w_k \ln p_k \varepsilon_{kj}\right) \\ &= \frac{\delta_{ij}}{w_i} - \frac{\beta_i}{w_i}\sum_k w_k \ln p_k \varepsilon_{kj} \end{aligned} \tag{3.A.8}$$

Each elasticity ε_{ij} is a function of itself and all other elasticities. Equation (3.A.8) can be expressed in matrix form and solved for elasticities using linear algebra.

$$E = A - BCE \tag{3.A.9}$$

where E is a (4 X 2) matrix containing elements $e_{ij} = \varepsilon_{ij}$, A is a (4 X 2) matrix with elements $a_{ij} = \frac{\delta_{ij}}{w_i}$, B is a (4 X 1) matrix with elements $b_i = \frac{\beta_i}{w_i}$ and C is a (1 X 4) matrix with elements $c_j = w_j \ln p_j$.

Equation (3.A.9) can be solved for E with the following steps:

$$
\begin{aligned}
&E + BCE = A \\
&(I + BC)E = A \\
&E = (I + BC)^{-1} A
\end{aligned}
\tag{3.A.10}
$$

Advertising elasticities for models with advertising expenditure as a price deflator (as in chapter 3.5.3) can be derived in a similar fashion. The final elasticity formula is

$$
\varepsilon_{ij} = -\frac{\gamma_{ij}\delta_j}{w_i} + \frac{\beta_i w_j \delta_j}{w_i} - \frac{\beta_i}{w_i}\sum_k w_k \left(\ln p_k - \delta_k \ln A_k\right)\varepsilon_{kj} \tag{3.A.11}
$$

which can be represented as

$$
E = -A + B - CDE \tag{3.A.12}
$$

E is a (4 X 2) matrix containing elements $e_{ij} = \varepsilon_{ij}$, A is a (4 X 2) matrix with elements $a_{ij} = \frac{\gamma_{ij}\delta_j}{w_i}$, B is a (4 X 2) matrix with elements $b_{ij} = \frac{\beta_i w_j \delta_j}{w_i}$, C is a (4 X 1) matrix containing elements $c_i = \frac{\beta_i}{w_i}$ and D is a (1 X 4) matrix with elements $d_j = w_j\left(\ln p_j - \delta_j \ln A_j\right)$. Note that the $\delta_j \ln A_j$ portion of the elements in D are zero for meats without advertising expenditure data (i.e. poultry and fish).

The solution to (A.12) is

$$
E = (I + CD)^{-1}(B - A) \tag{3.A.13}
$$

REFERENCES

"Livestock and Poultry Situation and Outlook Report." US Department of Agriculture. Economic Research Service.

Alston, J., and J. A. Chalfant. "Can We Take the Con Out of Meat Demand Studies?" *Western Journal of Agricultural Economics* 16, no. 1(1991): 36-48.

Alston, J., J. W. Freebairn, and J. S. James. "Beggar-Thy-Neighbour Advertising: Theory and Application to Generic Commodity Promotion Programs." *American Journal of Agricultural Economics* 83(2001): 888-902.

Anderson, G. J., and R. W. Blundell. "Estimation and Hypothesis Testing in Dynamic Singular Equation Systems." *Econometrica* 50, no. 6(1982): 1559-1571.

Anderson, G. J., and R. W. Blundell. "Consumer Non-Durable in the U.K.: A Dynamic System." *Economic Journal* 94(1984): 35-44.

Baye, M., D. Jansen, and J. Lee. "Advertising Effects in Complete Demand Systems." *Applied Economics* 24(1992): 1087-1096.

Boetel, B. L., and D. J. Liu. "Evaluating the Effect of Generic Advertising and Food Health Information Within a Meat Demand System." *Agribusiness* 19, no. 3(2003): 345-354.

Burton, M., and T. Young. "The impact of BSE on the demand for beef and other meats in Great Britain." *Applied Economics* 28(1996): 687-893.

Cavaliere, G., and G. Tassinari. "Advertising effect on primary demand: a cointegration approach." *International Journal of Advertising* 20(2001): 319-339.

Chang, H., and R. Green. 1992. "Measuring the Effects of Advertising on Demand Elasticities Using Time Series/Cross Sectional Data", in H. W. Kinnucan, S. R. Thompson, and H. Chang (ed.), Commodity Advertising and Promotion. Iowa State University Press, Ames.

Crespi, J. M. "The Generic Advertising Controversy: How Did We Get Here and Where Are We Going." *Review of Agricultural Economics* 25, no. 2(2003): 294-315.

Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *The American Economic Review* 70, no. 3(1980): 312-326.

Enders, W. *Applied Econometric Time Series*. Toronto: John Wiley & Sons Canada, 2004.

Engle, R. E., and C. W. F. Granger. "Cointegration and Error-Correction: Representation, Estimation and Testing." *Econometrica* 55(1987): 251-76.

Gordon, D. V., R. Hannesson, and W. A. Kerr. "What is a Commodity? An Empirical Definition Using Time Series Econometrics." *The Journal of International Food and Agribusiness Marketing* 10, no. 2(1999): 1-30.

Granger, C. W. F., and P. Newbold. "Spurious Regressions in Econometrics." *Journal of Econometrics* 2(1974): 111-120.

Green, R., and J. Alston. "Elasticities in AIDS Models." *American Journal of Agricultural Economics* 72(1990): 442-445.

Green, R., H. F. Carman, and K. McManus. "Some Empirical Methods of Estimating Advertising Effects in Demand Systems: An Application to Dried Fruits." *Western Journal of Agricultural Economics* 16, no. 1(1991): 63-71.

Greene, W.H. *Econometric Analysis*. Toronto: Prentice Hall of Canada Inc., 2002.

Hayes, D. 1992. "Discussion: Incorporating Advertising into Demand Systems", in H. W. Kinnucan, S. R. Thompson, and H. Chang (ed.), Commodity Advertising and Promotion. Iowa State University Press, Ames.

Johansen, S. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control* 12(1988): 231-254.

Kinnucan, H. W. "A note on measuring returns to generic advertising in interrelated markets." *Journal of Agricultural Economics* 47, no. 2(1996): 261-267.

Kinnucan, H. W., H. Chang, and M. Venkateswaran. "Generic Advertising Wearout." *Review of Marketing and Agricultural Economics* 61, no. 3(1993): 401-415.

Kinnucan, H. W., S. R. Thompson, and H. Chang. *Commodity Advertising and Promotion*. Edited by H. W. Kinnucan, S. R. Thompson, and H. Chang. Ames: Iowa State University Press, 1992.

Leading National Advertisers. *Ad $ Summary*. New York, all issues 1975-2001.

Lenz, J., H. M. Kaiser, and C. Chung. "Economic Analysis of Generic Milk Advertising Impacts on Markets in New York State." *Agribusiness* (1998): 73-83.

MacDonald, A., and P. Gould. 1992. "Generic Advertising: A Commodity Perspective", in H. W. Kinnucan, S. R. Thompson, and H. Chang (ed.), Commodity Advertising and Promotion. Iowa State University Press, Ames.

Piggott, N. E., et al. "Demand Response to Advertising in the Australian Meat Industry." *American Journal of Agricultural Economics* 78(1996): 268-279.

Puttman, J. J., and J. E. Allshouse. "Food consumption, price and expenditure, 1970-1997." US Department of Agriculture, ERS Statistical Bulletin Number 840.

Reberte, J., T. Schmit, and H. M. Kaiser. "An Economic Analysis of Generic Egg Advertising in California, 1985-1995." *Agribusiness* 13(1997): 365-373.

Rickertsen, K., J. A. Chalfant, and M. Steen. "The effects of advertising on the demand for vegetables." *European Review of Agricultural Economics* 22(1995): 481-494.

Stone, F. R. N. *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom*. Vol. 1. Cambridge: Cambridge Press, 1953.

Thiel, H. *System-Wide Explorations in International Economics, Input-Output Analysis, and Marketing Research*. New York: North Holland Publishing Company, 1980.

US Department of Agriculture. Economic Research Service. *Livestock and Poultry Situation and Outlook Report*. All issues 1976-2002.

Zellner, A. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American Statistical Association* (1962): 348-368.